



Optimal Portfolio with Vector Expected Utility

Eric André

► To cite this version:

| Eric André. Optimal Portfolio with Vector Expected Utility. 2013. halshs-00796482

HAL Id: halshs-00796482

<https://shs.hal.science/halshs-00796482>

Preprint submitted on 4 Mar 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Optimal Portfolio with Vector Expected Utility

Eric André

WP 2013 - Nr 08

Optimal Portfolio with Vector Expected Utility

Eric ANDRÉ*

Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS

February 11, 2013

Abstract

We study the optimal portfolio selected by an investor who conforms to Siniscalchi (2009)'s Vector Expected Utility's (VEU) axioms and who is ambiguity averse. To this end, we derive a mean-variance preference generalised to ambiguity from the second-order Taylor-Young expansion of the VEU certainty equivalent. We apply this Mean Variance Variability preference to the static two-assets portfolio problem and deduce asset allocation results which extend the mean-variance analysis to ambiguity in the VEU framework. Our criterion has attractive features: it is axiomatically well-founded and analytically tractable, it is therefore well suited for applications to asset pricing as proved by a novel analysis of the home-bias puzzle with two ambiguous assets.

JEL classification: D81, G11.

Keywords: Vector Expected Utility, Ambiguity, Portfolio Choice, Home-bias Puzzle.

1 Introduction

Since the seminal works of Markowitz (1952) and Tobin (1958), mean-variance preferences have been the cornerstone of optimal portfolio theory. An investor with mean variance preference having to select risky assets will rank uncertain portfolio returns r according to the following evaluation of their utility:

$$u_{MV}(r) = \mathbf{E}_{\mathbf{p}}(r) - \frac{\gamma}{2} \mathbf{var}_{\mathbf{p}}(r)$$

where \mathbf{p} is a given probability and γ is a measure of the aversion to variance. The foundations of the mean-variance preferences and the link between risk and variance for “small risks” are to be found in the classical Arrow-Pratt (Pratt, 1964; Arrow, 1965) approximation of the Expected Utility (EU) certainty equivalent:

*I thank André Lapiéd for guidance and comments that helped to substantially improve the paper. E-mail: eric.andre@univ-amu.fr. Address: AMSE-GREQAM, 3 Avenue Robert Schuman, 13628 Aix-en-Provence Cedex 01.

for an investor with certain wealth w considering a risky investment h , the Taylor expansion to the second order of its certainty equivalent is given by:

$$c(w + h) = w + \mathbf{E}_{\mathbf{p}}(h) - \frac{1}{2}\gamma(w) \mathbf{var}_{\mathbf{p}}(h) + o(\mathbf{var}_{\mathbf{p}}(h))$$

where $\gamma(w) = -u''(w)/u'(w)$ is the absolute risk aversion coefficient of the *Bernoulli utility function*¹ u .

While the mean-variance analysis remains the workhorse of modern portfolio theory, it is well known that empirical data cannot be fully rationalised in this context, especially, the equity premium cannot be explained by a risk premium only (the “equity premium puzzle”, [Mehra and Prescott, 1985](#)) and international portfolios are under diversified (the “home-bias puzzle”, [French and Poterba, 1991](#)). A large literature has endeavoured to explain these “puzzles”, analysing different shortcomings of the classical paradigm. Among these, recent advances in decision theory aimed at generalising the EU framework have allowed to study the effect on asset prices of *ambiguity*: situations where the information available to the investor is too imprecise to be summarised by a unique probability distribution over events. This paper fits into this field of research: its main contribution is to propose a mean-variance preference generalised to ambiguity using [Siniscalchi \(2009\)](#)’s Vector Expected Utility (VEU). We study the conditions for existence and calculate the second order Taylor-Young expansion of the VEU certainty equivalent from which we derive a *Mean Variance Variability* preference. This flexible and tractable criterion allows not only to retrieve the existing results for an optimal portfolio with one risky and one ambiguous asset but also to show new results with two ambiguous assets, which we apply to the discussion of the home-bias puzzle.

Several non EU decision theoretic models have been successfully applied to the field of finance and to the discussions of the “puzzles”. Among these applications, some have sought to improve the mean-variance preferences: especially [Maccheroni, Marinacci, and Ruffino \(Forthcoming\)](#) derive a mean-variance model adjusted for ambiguity from a quadratic approximation of the certainty equivalent of the smooth model of decision making under ambiguity ([Klibanoff, Marinacci, and Mukerji, 2005](#), henceforth [KMM](#)). Our work is closely related to this paper which provided the impetus for our research, but is set in a different axiomatic framework hence uses different mathematical tools: while [KMM](#) introduce second-order acts and assume subjective expected utility over them, [Siniscalchi \(2009\)](#) derives the VEU representation formula from behavioural axioms. Moreover, a baseline probability has to be hypothesised in the smooth model while it is endogenously derived from the preferences of the decision maker in the VEU model as will be detailed below.

A decision maker (DM) conforms to the VEU set of axioms if and only if (iff) she ranks uncertain prospects f , functions from a state space S to a consequence space X , via the functional:

$$V(f) = \mathbf{E}_{\mathbf{p}}(\tilde{u} \circ f) + A(\{\mathbf{E}_{\mathbf{p}}(\zeta_i \cdot \tilde{u} \circ f)\}_{0 \leq i < n}) \quad (1)$$

where $\tilde{u}: X \rightarrow \mathbb{R}$ is a von Neumann-Morgenstern utility function and $A: \mathbb{R}^n \rightarrow \mathbb{R}$ is a function such that for any $d \in \mathbb{R}^n$, $A(-d) = A(d)$ and $A(0_n) = 0$. There

¹In [Mas-Colell, Whinston, and Green \(1995\)](#)’s terminology, see note 12 p. 184.

are two parts to this evaluation: the expected utility of the act according to a baseline probability \mathbf{p} and an adjustment to this baseline evaluation, function of the *variability* of the utility profile and of the DM's attitude toward ambiguity.

The baseline probability \mathbf{p} is a key feature of the VEU model as it is revealed by the preferences of the DM over *complementary acts*. These are acts f and \bar{f} such that for any states s and $s' \in S$, $\frac{1}{2}f(s) + \frac{1}{2}\bar{f}(s) \sim \frac{1}{2}f(s') + \frac{1}{2}\bar{f}(s')$ which implies that their utility profiles sum to a constant: $\tilde{u} \circ f = k - \tilde{u} \circ \bar{f}$ for some $k \in \mathbb{R}$. In the words of Siniscalchi, they are “the preference counterpart of algebraic negation”. In a portfolio application, assuming linear utility, a long and a short position of the same value in the same asset are straightforward examples of complementary acts. The central insight of the VEU model is that complementary acts have the same utility variability, *i.e.* the same ambiguity, hence have to be ranked according to their baseline expected utility only. Therefore “preferences over complementary acts uniquely identify the baseline prior”. As an illustration, consider Ellsberg (1961)'s three colour single urn experiment: a ball is drawn from an urn containing 30 red balls and 60 black and yellow balls with the proportion of black and yellow balls unknown. Assuming linear utility, the act $(10, R; -10, B; 0, Y)$ that yields \$10 if a red ball is drawn and costs \$10 if a black ball is drawn and the act $(-10, R; 10, B; 0, Y)$ are complementary: they embed the same ambiguity and if the DM is indifferent between these two, we can derive that $\mathbf{p}(R) = \mathbf{p}(B)$. If the DM is also indifferent between $(10, R; 0, B; -10, Y)$ and $(-10, R; 0, B; 10, Y)$, we can infer that she is using the uniform prior as her baseline probability over the state space $S = \{R, B, Y\}$.

The adjustment to the baseline evaluation refer to the notion of *crisp* acts, which has been introduced by Ghirardato, Maccheroni, and Marinacci (2004, henceforth GMM). They characterise the *unambiguous preference* as the maximal² restriction satisfying the independence axiom of the complete DM preference. This preference is incomplete and has a Bewley (2002) representation by a unanimity criterion over a set of priors \mathcal{C} . Acts which have the same expected utility for all the priors in \mathcal{C} are crisps, hence non crisps acts have “variable utility profiles”. Siniscalchi (2009) proves that the subspace of crisp acts \mathbf{C} and the subspace of non crisp acts \mathbf{NC} are orthogonal complements in $\mathbf{L}^2(\mathbf{p})$ and constructs an orthonormal basis of \mathbf{NC} : the family $\{\zeta_i\}_{0 \leq i < n}$ of adjustment factors which can be interpreted as independent sources of ambiguity. Going back to equation (1), it can now be seen that the vectorial argument of the adjustment function A is the vector of coordinates of the utility profile in \mathbf{NC} , which can be read as the correlations of the utility profile with each source of ambiguity. Thanks to this construction the VEU evaluation nicely reduces to EU for crisp acts and reflects complementarities among ambiguous acts. This can again be illustrated using Ellsberg's three colour urn, as in the original paper: let ζ_0 be the random variable such that $\zeta_0(R) = 0$, $\zeta_0(B) = 1$ and $\zeta_0(Y) = -1$ and let $A(\phi) = -|\phi|$. Assuming the uniform prior that we derived above, any act f is evaluated through: $V(f) = \frac{1}{3}(f(R) + f(B) + f(Y)) - |\frac{1}{3}(f(B) - f(Y))|$. One can check that this evaluation is consistent with the preferences reported in Ellsberg (1961): $V[(10, R; 0, B; 0, Y)] > V[(0, R; 10, B; 0, Y)]$ but $V[(10, R; 0, B; 10, Y)] < V[(0, R; 10, B; 10, Y)]$ highlighting the complementarity of the payoffs on the events B and Y in the last act.

²In the sense of set inclusion.

The mathematical details of this setup are exposed in Section 2. In Section 3 the conditions under which the first and second order differentials of the VEU certainty equivalent exist are proven and the calculus of the second-order Taylor-Young expansion is detailed. This quadratic approximation for a “small” incremental act is first exposed in the most general case, then in the case where the DM is *ambiguity averse* in the sense of [Schmeidler \(1989\)](#), *i.e.* a DM with a weak preference for mixtures. The ambiguity averse case highlights the central role of the Hessian of the adjustment function A : the \mathbb{R}^n rotation matrix made of its eigenvectors induces a rotated basis of the subspace \mathbf{NC} of non-crisp acts made of normalised sources of ambiguity which are still \mathbf{p} -independent but are also not correlated for the DM’s tastes in the sense that there are no cross terms in her evaluation of an act spanning several sources. Section 4 delves further into the properties of the quadratic approximation and proves the link between the DM aversion to these sources of ambiguity and the eigenvalues of the Hessian of A . The ambiguity adjustment term in the quadratic approximation can then be written as the variance of a purely non crisp act $\mathbb{A}h$ obtained by scaling and projecting the evaluated prospect h . Building on this analysis, Section 4 proceeds with the exposition of our generalised Mean Variance Variability criterion which takes the form:

$$u_{\text{MVV}}(h) = \mathbf{E}_{\mathbf{p}}(h) - \frac{\gamma}{2} \mathbf{var}_{\mathbf{p}}(h) - \frac{\theta}{2} \mathbf{var}_{\mathbf{p}}(\mathbb{A}h).$$

where γ and θ are strictly positive constant that respectively measure risk aversion and ambiguity aversion. Section 4 concludes with the similarities and differences between our criterion and [Maccheroni et al. \(Forthcoming\)](#)’s. Section 5 applies the Mean Variance Variability to the choice of an optimal portfolio and shows that in our setting, the ambiguity adjustment leads to a modified variance-covariance matrix for the assets, a results that lends a lot of tractability to our criterion. Finally an analysis of the home-bias puzzle is proposed, first with a purely risky domestic asset and an ambiguous foreign asset, then with two ambiguous assets, a new analysis allowed by our criterion. Proofs are grouped in the appendix.

Related literature. As has been highlighted above, our work is strongly related to [Maccheroni et al. \(Forthcoming\)](#) which also contains a study of ambiguity through a second-order approximation “in the small”, a mean-variance analysis of the optimal portfolio generalised to ambiguity and an application to the discussion of the home-bias puzzle. Other works share at least one of these lines of research.

[Nau \(2003, 2006\)](#), [Skiadas \(2008\)](#), [Izhakian and Benninga \(2011\)](#) and [Jewitt and Mukerji \(2011\)](#) study ambiguity in the small. [Nau \(2003\)](#) studies the second order approximation of the risk premium for non EU preferences in the state-preference framework and obtains a generalised measure of risk aversion given by the Hessian matrix of the risk neutral probability distribution, where the variables are the finite states of the world. [Nau \(2006\)](#) applies these results to the axiomatisation of a second order utility preference and exhibits a decomposition of the second order approximation of the risk premium into a “pure risk” and an “uncertainty” premium. [Skiadas \(2008\)](#) studies approximations of KMM

and divergence³ preferences' certainty equivalents, for small risks of Brownian and Poissonian types. [Izhakian and Benninga \(2011\)](#) compute in the KMM framework an approximated uncertainty premium with an ambiguity premium component. [Jewitt and Mukerji \(2011\)](#), building on their definitions of “more ambiguous” acts, expose a second order approximation of the ambiguity premium in the KMM and the α -MEU (see e.g. [GMM](#)) frameworks. By extending this analysis to the VEU model, our paper allows to test the robustness of these results to the axiomatic specifications.

[Taboga \(2005\)](#), [Maccheroni, Marinacci, Rustichini, and Taboga \(2009\)](#), [Garlappi, Uppal, and Wang \(2007\)](#) and [Boyle, Garlappi, Uppal, and Wang \(2012\)](#) apply a mean variance generalised to ambiguity to the analysis of the optimal portfolio. [Taboga \(2005\)](#) shows that in the KMM model, with Constant Absolute Risk Aversion (CARA) and Constant Ambiguity Attitude, [Markowitz \(1952\)](#) results hold if unique values of expected returns, variances and covariances are replaced by averages of these moments calculated for all the distributions of asset returns considered by the DM. Furthermore, these averages are not computed under KMM preference's second order probability but under a modified law that puts more weights on pessimistic priors. [Maccheroni et al. \(2009\)](#) build from the variational model mean-variance preferences which are monotone. Applied to the optimal portfolio, they lead to another generalisation of [Markowitz'](#) results where the unconditional means and variances are replaced by moments conditioned on the wealth not exceeding a given threshold: the investor ignores the most favourable parts of the distributions, parts which lead to high variances hence to the non monotonicity of the original mean-variance preferences. [Garlappi et al. \(2007\)](#) study the effect of uncertainty on estimated parameters when the DM uses a Maxmin Expected Utility ([Gilboa and Schmeidler, 1989](#), henceforth MEU) approach versus a Bayesian approach and apply their results to the optimal portfolio. [Boyle et al. \(2012\)](#) propose a mean variance analysis with confidence intervals on the means of the returns and an MEU investor. They show that with increasing ambiguity, *i.e.* increasing widths of the confidence intervals, the optimal portfolio changes from [Markowitz'](#) diversified holdings to concentration on familiar assets to non participation. Our results confirm that [Markowitz'](#) results hold in another axiomatic decision making model if classical variances and covariances are adjusted for ambiguity, these adjustments being model specific.

Our work is also related to several papers which explore the effect of ambiguity on asset prices and possible implications for the puzzles. Among others, [Chen and Epstein \(2002\)](#) study asset returns with a continuous time multiple priors setting and obtain a decomposition of the equity premium in terms of risk and ambiguity, possibly covering some of the gap between historical data and the purely risky equity premium. [Epstein and Miao \(2003\)](#) use a recursive multiple prior settings to characterise the equilibrium in an economy with two agents with heterogeneous beliefs and apply this result to the discussion of the home-bias puzzle. Finally [Gollier \(2011\)](#) studies the comparative statics of more ambiguity aversion and exposes the condition for an increase in ambiguity aversion to decrease the ambiguous asset holding.

³[Maccheroni, Marinacci, and Rustichini \(2006\)](#) variational preferences with the Fenchel conjugate equal to the divergence of the priors with respect to a baseline probability.

2 Setup and notations

We use the [Anscombe and Aumann \(1963, AA\)](#) setup where acts f are functions from a state space S , endowed with a countably generated⁴ σ -algebra of events Σ , to the set of all simple probability distributions over an arbitrary set of prizes Z . We denote by \mathcal{F} the set of all AA acts.

We consider a DM who satisfies [Siniscalchi \(2009\)](#)'s VEU's axioms. Therefore her preferences are represented by a sharp VEU functional $(\tilde{u}, \mathbf{p}, n, \{\zeta_i\}_{0 \leq i < n}, A)$ such that $f \succsim g$ iff $V(f) \geq V(g)$ where:

$$V(f) = \mathbf{E}_{\mathbf{p}}(\tilde{u} \circ f) + A(\{\mathbf{E}_{\mathbf{p}}(\zeta_i \cdot \tilde{u} \circ f)\}_{0 \leq i < n}) \quad (2)$$

with $\tilde{u} \circ f(s) = \mathbf{E}_{f(s)}[u] = \int_Z u(z) df(s)(z)$ for all $s \in S$.

The function $A: \mathbb{R}^n \rightarrow \mathbb{R}$ is such that for any $d \in \mathbb{R}^n$, $A(-d) = A(d)$ and $A(0_n) = 0$ and u is the Bernoulli utility function on which we make the usual assumptions:

Assumption 1. *u is of class \mathcal{C}^2 , strictly increasing and concave.*

Recall that \mathbf{p} is a baseline probability over the measured space (S, Σ) which is revealed by the DM preferences over complementary acts. The set \mathcal{C} of priors given by the [Bewley \(2002\)](#) representation of the incomplete unambiguous preference (see [GMM](#)) is symmetric around this baseline probability and all the probabilities in this set are absolutely continuous with respect to \mathbf{p} ([Siniscalchi, 2009](#), Lemma 3).

An act k is *crisp*⁵ if and only if it behaves like its certainty equivalent that is for all $x \in \mathbb{R}$ such that $k \sim x\mathbf{1}_S$, for all AA act g , for all $\alpha \in (0, 1]$ we have $\alpha g + (1 - \alpha)x\mathbf{1}_S \sim \alpha g + (1 - \alpha)k$. A remarkable property of crisp acts is that their expected utility is the same for all the probabilities in the set of priors \mathcal{C} . Having no variation of its utility profile, the vector φ associated to a crisp k is 0_n hence its adjustment is null and its VEU evaluation reduces to $V(k) = \mathbf{E}_{\mathbf{p}}(\mathbf{E}_{k(s)}[u])$.

The family $\zeta = \{\zeta_i\}_{0 \leq i < n}$ is the orthonormal basis in $L^2(\mathbf{p})$ of the subspace \mathbf{NC} of non crisp acts, whose orthogonal complement is the subspace \mathbf{C} of crisps acts. While the VEU model allows for a countably infinite family, we will make the assumption in the following that n is finite. The number of adjustment factors is discussed in [Siniscalchi \(2009, section 4.2.\)](#) and its Proposition 1 implies that, in an asset pricing application, there exists a crisp portfolio for any m assets when $m > n$.

As we are interested in monetary outcomes, we set Z to be the real line. Following [Kreps \(1988\)](#), we identify the set of our objects of choice, random variables $X: S \rightarrow \mathbb{R}$, with a subset denoted \mathcal{F}_d of the AA acts: those whose values are the degenerate lotteries $\delta_{X(s)}$ for each $s \in S$. Therefore we have $\tilde{u} \circ f(s) = \mathbf{E}_{f(s)}[u] = u(X(s))$ hence we will denote by f either the act in \mathcal{F}_d or the associated random variable.

As all the relevant probabilities for the DM are absolutely continuous with respect to \mathbf{p} , she will not discriminate between two acts which are equal \mathbf{p} -almost

⁴For a discussion of this assumption, see [Siniscalchi \(2009, Section 2\)](#).

⁵We use [Siniscalchi \(2009\)](#)'s definition, see his paper for differences with [GMM](#)'s.

everywhere. This allows us to embed the space \mathcal{F}_d into $L^\infty(\mathbf{p})$, the Banach space of equivalent classes of essentially bounded Σ -measurable functions. As there is only one underlying measured space, we'll drop the reference to \mathbf{p} and use L^∞ , the associated norm will be denoted by $\|\cdot\|_\infty$.

We introduce three self-explanatory additional functions from L^∞ to L^∞ that will be used in the remainder of this paper: $U: f \mapsto u \circ f$, $U': f \mapsto u' \circ f$ and $U'': f \mapsto u'' \circ f$, and, slightly changing the initial notations for the VEU model, we define the family of functions $\{\varphi_i\}_{0 \leq i < n}$ from L^∞ to \mathbb{R} such that $\varphi_i(a) = \mathbf{E}_{\mathbf{p}}(\zeta_i \cdot a)$ and the associated vector valued function $\varphi: L^\infty \rightarrow \mathbb{R}^n$ such that $\varphi(a) = [\varphi_0(a), \dots, \varphi_{n-1}(a)]^\top$. With these additional notations, we can rewrite equation (2) for any $f \in \mathcal{F}_d$ as:

$$V(f) = \mathbf{E}_{\mathbf{p}}(U(f)) + A(\varphi(U(f))) \quad (3)$$

3 Quadratic approximation of the certainty equivalent

The certainty equivalent⁶ is a function c from L^∞ to \mathbb{R} such that for any act f we have $c(f)\mathbf{1}_S \sim f$ which is $V(c(f)\mathbf{1}_S) = u(c(f)) = V(f)$ and finally $c(f) = u^{-1} \circ V(f)$.

We want to write the second order Taylor-Young expansion of the certainty equivalent around an act w (the initial portfolio or initial wealth) in the direction of a “small” incremental act h :

$$c(w + h) = c(w) + d_w c(h) + \frac{1}{2} d_w^2 c(h)^2 + o(\|h\|_\infty^2) \quad (4)$$

where $d_w c: L^\infty \rightarrow \mathbb{R}$ is the differential at point w . We will also write $dc: L^\infty \rightarrow \mathcal{L}(L^\infty, \mathbb{R})$ for the differential function and $d^2 c = d(dc): L^\infty \rightarrow \mathcal{L}(L^\infty, \mathcal{L}(L^\infty, \mathbb{R}))$ for the second order differential function.

For these differentials to exist, we need to make the following assumption:

Assumption 2. *A is twice continuously differentiable over \mathbb{R}^n .*

We will denote by $\nabla A(x)$ the gradient vector of the function A at point $x \in \mathbb{R}^n$, by $\nabla^2 A(x)$ its Hessian matrix at point x , by $\langle x, y \rangle$ and $\|x\|_{\mathbb{R}^n}$ the scalar product and norm in \mathbb{R}^n . Given that φ is a linear mapping (hence is of class \mathcal{C}^∞), assumptions 1 and 2 and the following proposition imply that c is twice differentiable (proofs are in the appendix).

Proposition 3.1. *U is twice L^∞ -differentiable and we have for all $a, h, k \in L^\infty$: $d_a U(h) = hU'(a)$, $d_a^2 U(h, k) = hkU''(a)$.*

3.1 Quadratic approximation in the general case

We can now state the results, where $\gamma(c(w)) := -u''(c(w))/u'(c(w))$ is the Arrow-Pratt coefficient of risk aversion valued at the certainty equivalent of the initial wealth, $\hat{w} := \varphi(U(w))$ is the \mathbb{R}^n vector of coordinates of the projection onto the subspace \mathbf{NC} of the initial wealth's utility profile and $\dot{w} := U'(w)/u'(c(w))$ and $\ddot{w} := U''(w)/u''(c(w))$ are functions in L^∞ .

⁶Its existence and uniqueness is a consequence of the continuity and strict monotonicity of u over the connected space \mathbb{R} .

Proposition 3.2. *The first order differential at point $w \in \mathbf{L}^\infty$ in direction $h \in \mathbf{L}^\infty$ is given by :*

$$d_w c(h) = \mathbf{E}_{\mathbf{P}}(h\dot{w}) + \langle \nabla A(\hat{w}), \varphi(h\dot{w}) \rangle \quad (5)$$

Proposition 3.3. *The second order differential at point $w \in \mathbf{L}^\infty$ in direction $h \in \mathbf{L}^\infty$ valued at h is given by :*

$$\begin{aligned} d_w^2 c(h)^2 = & -\gamma(c(w))(\mathbf{E}_{\mathbf{P}}(h^2\ddot{w}) - (\mathbf{E}_{\mathbf{P}}(h\dot{w}))^2) \\ & - \gamma(c(w))(\langle \nabla A(\hat{w}), \varphi(h^2\ddot{w}) \rangle - \langle \nabla A(\hat{w}), \varphi(h\dot{w}) \rangle^2) \\ & + 2\gamma(c(w))\mathbf{E}_{\mathbf{P}}(h\dot{w})\langle \nabla A(\hat{w}), \varphi(h\dot{w}) \rangle \\ & + u'(c(w))\langle \nabla^2 A(\hat{w})\varphi(h\dot{w}), \varphi(h\dot{w}) \rangle \end{aligned} \quad (6)$$

Special cases. The above formulae lead to some well known results when additional assumptions are made.

If the act w is the current wealth, that is w is the degenerate random variable $w\mathbf{1}_{\mathbf{S}}$ where we also denote by w the scalar measuring this current wealth, then $c(w\mathbf{1}_{\mathbf{S}}) = w$, $\hat{w} = 0_n$, and $\dot{w} = \ddot{w} = \mathbf{1}_{\mathbf{S}}$. The second order Taylor-Young expansion is given by:

$$\begin{aligned} c(w + h) = & w + \mathbf{E}_{\mathbf{P}}(h) + \langle \nabla A(0_n), \varphi(h) \rangle - \frac{1}{2}\gamma(w)\mathbf{var}_{\mathbf{P}}(h) \\ & - \frac{1}{2}\gamma(w)(\langle \nabla A(0_n), \varphi(h^2) \rangle - \langle \nabla A(0_n), \varphi(h) \rangle^2) \\ & + \gamma(w)\mathbf{E}_{\mathbf{P}}(h)\langle \nabla A(0_n), \varphi(h) \rangle \\ & + \frac{1}{2}u'(w)\langle \nabla^2 A(0_n)\varphi(h), \varphi(h) \rangle + o(\|h\|_\infty^2). \end{aligned} \quad (7)$$

If w is the current wealth and h is crisp, then $\varphi(h) = 0_n$ and the certainty equivalent reduces, as expected, to the Arrow-Pratt approximation :

$$c(w + h) = w + \mathbf{E}_{\mathbf{P}}(h) - \frac{1}{2}\gamma(w)\mathbf{var}_{\mathbf{P}}(h) + o(\|h\|_\infty^2) \quad (8)$$

In our paper, we will consider the case, detailed in the following subsection, where w is the current wealth, h is not necessarily crisp and the DM is ambiguity averse.

3.2 Quadratic approximation with ambiguity aversion

We will suppose in the remaining of the paper that the DM conforms to the Ambiguity Aversion axiom (Axiom 9 in [Siniscalchi, 2009](#)), that is for all AA acts f and g and for all $\alpha \in (0, 1)$, $f \sim g$ implies $\alpha f + (1 - \alpha)g \succsim g$. [Siniscalchi \(2009, Corollary 2\)](#) proves that this is equivalent to A being nonpositive and concave.

Note that with this axiom, we could obtain Propositions 3.2 and 3.3 without Assumption 2, indeed [Rockafellar \(2000, Theorem 2.8 and Corollary 2.9\)](#) proves the existence almost everywhere of the second-order expansion of a closed, proper, convex function $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$. Nonetheless, this would not ensure that A is differentiable at 0_n and that we can write equation (7).

As A is symmetric, *ie* $A(-\varphi) = A(\varphi)$, and $A(0_n) = 0$, the concavity implies that A reaches its maximum value 0 at point 0_n that is $\nabla A(0_n) = 0_n$ (our smoothness assumptions can be related to [Segal and Spivak \(1990\)](#)'s attitude toward risk of order 2). With the Ambiguity Aversion axiom equation (7) reduces to:

$$c(w + h) = w + \mathbf{E}_{\mathbf{p}}(h) - \frac{1}{2}\gamma(w) \mathbf{var}_{\mathbf{p}}(h) + \frac{1}{2}u'(w) \langle \nabla^2 A(0_n) \cdot \varphi(h), \varphi(h) \rangle + o(\|h\|_{\infty}^2) \quad (9)$$

Being a Hessian matrix, $\nabla^2 A(0_n)$ is symmetric and A being concave, it is semi-definite negative ([Rockafellar, 1970](#), Theorem 4.5).

Diagonalisation of the Hessian. The symmetry of $\nabla^2 A(0_n)$ implies that there exists an orthonormal basis $\{\hat{e}_i\}_{0 \leq i < n}$ of \mathbb{R}^n formed with normalised eigenvectors ([Golub and Van Loan, 1996](#), Theorem 8.1.1). Assume w.l.o.g. that the eigenvalues are ordered: $-\lambda_0 \geq \dots \geq -\lambda_{n-1} \geq 0$, there exists a real orthogonal matrix $Q = (\hat{e}_0, \dots, \hat{e}_{n-1})$ such that $Q^T \nabla^2 A(0_n) Q = \text{diag}(\lambda_0, \dots, \lambda_{n-1})$ where λ_i is the eigenvalue associated with the eigenvector \hat{e}_i . We will denote by Λ the diagonal matrix with positive elements $\Lambda := \text{diag}(-\lambda_0, \dots, -\lambda_{n-1})$.

Q is a rotation matrix (of generic element q_{ij}) and, $\{e_i\}_{0 \leq i < n}$ being the canonical basis of \mathbb{R}^n , the vectors $\hat{e}_i = Q e_i$ define the rotated basis of \mathbb{R}^n in which the Hessian matrix is diagonal. Now φ is an isomorphism from \mathbf{NC} (which we've supposed to be finite dimensional) to \mathbb{R}^n mapping the basis $\zeta = \{\zeta_i\}_{0 \leq i < n}$ to $\{e_i\}_{0 \leq i < n}$. Define the family $\hat{\zeta} = \{\hat{\zeta}_i\}_{0 \leq i < n}$ by $\hat{\zeta}_i = \sum_{j=0}^{n-1} q_{ji} \zeta_j$. We have that

$$\mathbf{E}_{\mathbf{p}}(\hat{\zeta}_i \cdot \hat{\zeta}_j) = \mathbf{E}_{\mathbf{p}} \left(\left(\sum_{k=0}^{n-1} q_{ki} \zeta_k \right) \left(\sum_{l=0}^{n-1} q_{lj} \zeta_l \right) \right) = \sum_{k=0}^{n-1} \sum_{l=0}^{n-1} q_{ki} q_{lj} \mathbf{E}_{\mathbf{p}}(\zeta_k \cdot \zeta_l).$$

As ζ is an orthonormal basis, $\mathbf{E}_{\mathbf{p}}(\zeta_k \cdot \zeta_l) = \delta_{kl}$ (Kronecker delta) and $\mathbf{E}_{\mathbf{p}}(\hat{\zeta}_i \cdot \hat{\zeta}_j) = \sum_{k=0}^{n-1} q_{ki} q_{kj}$ which is the expression of the element (i, j) of the product of matrices $Q^T Q = \text{Id}$, hence $\mathbf{E}_{\mathbf{p}}(\hat{\zeta}_i \cdot \hat{\zeta}_j) = \delta_{ij}$ and $\hat{\zeta}$ is an orthonormal basis of \mathbf{NC} . Finally, the mapping $\hat{\varphi}$ defined by $\hat{\varphi}_i(a) = \mathbf{E}_{\mathbf{p}}(\hat{\zeta}_i \cdot a)$ is an isomorphism from \mathbf{NC} to \mathbb{R}^n mapping the basis $\hat{\zeta}$ to $\{\hat{e}_i\}_{0 \leq i < n}$.

Taking the coordinates of the projection of h in this rotated basis $\hat{\zeta}$ of the space \mathbf{NC} allows to use the diagonalised Hessian of A . Then we can rewrite equation (9) as:

$$c(w + h) = w + \mathbf{E}_{\mathbf{p}}(h) - \frac{1}{2}\gamma(w) \mathbf{var}_{\mathbf{p}}(h) - \frac{1}{2}u'(w) \langle \Lambda \cdot \hat{\varphi}(h), \hat{\varphi}(h) \rangle + o(\|h\|_{\infty}^2) \quad (10)$$

Interpretation. There is a direct interpretation of equation (10). Recall that the VEU criterion consists of two parts: an EU evaluation using the baseline probability \mathbf{p} and an adjustment, non positive and concave when the DM is ambiguity averse, function of the variability of the utility profile of the act considered. The quadratic approximation also consists of two parts: the first three terms are the

classical Arrow-Pratt approximation of the EU certainty equivalent and the last term is, to a factor $u'(w)$, the approximation of the adjustment for the variability of the utility profile. Indeed the Taylor-Young approximation to the second order of A at 0_n in direction $d \in \mathbb{R}^n$ (endowed with the basis of eigenvectors of the Hessian of A) is:

$$A(d) = A(0_n) + \langle \nabla A(0_n), d \rangle - \frac{1}{2} \langle \Lambda d, d \rangle + o(\|d\|_{\mathbb{R}^n}^2)$$

which, when the DM is ambiguity averse, reduces to:

$$A(d) = -\frac{1}{2} \langle \Lambda d, d \rangle + o(\|d\|_{\mathbb{R}^n}^2).$$

Now, if we assume that the Bernoulli utility function can be locally approximated around the point w by an affine function that is we can write $U(w\mathbf{1}_S + h) \approx u(w) + u'(w)h$, then we get that $\hat{\varphi}(U(w\mathbf{1}_S + h)) \approx u'(w)\hat{\varphi}(h)$ and:

$$A(\hat{\varphi}(U(w\mathbf{1}_S + h))) = -\frac{1}{2}(u'(w))^2 \langle \Lambda \hat{\varphi}(h), \hat{\varphi}(h) \rangle + o(\|\hat{\varphi}(h)\|_{\mathbb{R}^n}^2).$$

4 Mean Variance Variability Preference

4.1 Further analysis of the quadratic approximation in the ambiguity averse case

The ambiguity adjustment given by the fourth term of equation (10)

$$-\frac{1}{2}u'(w) \langle \Lambda \cdot \hat{\varphi}(h), \hat{\varphi}(h) \rangle \quad (11)$$

has two components: the variability of the utility profile of h given by $\hat{\varphi}(h)$, and the attitude toward ambiguity of the DM given by Λ . The later has itself two components: the DM overall level of aversion to ambiguity and her specific levels of aversion to each sources of ambiguity. We would like to have a measure of the respective contributions of this three effects using vector and matrix norms.

Starting with the overall level of aversion to ambiguity, the following property clarifies the link between the Hessian matrix and ambiguity aversion ranking as defined by [Ghirardato and Marinacci \(2002\)](#). Recall that they define⁷ a preference \succsim_1 to be *more ambiguity averse* than a preference \succsim_2 , if and only if for all $f \in \mathcal{F}_d$ and $x \in X$, $f \succsim_1 x \Rightarrow f \succsim_2 x$. [Siniscalchi \(2009, Proposition 4\)](#) proves that for two VEU preferences $(u, \mathbf{p}, n, \zeta, A^1)$ and $(u, \mathbf{p}, n, \zeta, A^2)$ this is equivalent to $A^1(d) \leq A^2(d)$ for all $d \in \mathbb{R}^n$. In our setting this extends to:

Proposition 4.1. *Consider two VEU preferences \succsim_1 and \succsim_2 on \mathcal{F}_d satisfying the Ambiguity Aversion Axiom with (sharp, Hessian adjusted) representations $(u, \mathbf{p}, n, \hat{\zeta}, A^1)$ and $(u, \mathbf{p}, n, \hat{\zeta}, A^2)$ where A^1 and A^2 are twice continuously differentiable with respective ordered diagonal Hessian matrices $\Lambda^1 = \text{diag}(\lambda_i^1)$ and $\Lambda^2 = \text{diag}(\lambda_i^2)$. \succsim_1 is more ambiguity averse than \succsim_2 if and only if for all $0 \leq i < n$:*

$$|\lambda_i^1| \geq |\lambda_i^2|.$$

⁷This is the version in [Siniscalchi \(2009, Definition 4\)](#).

Then if \succsim_1 is more ambiguity averse than \succsim_2 , any matrix norm will be greater for $\nabla^2 A^1(0_n)$ than for $\nabla^2 A^2(0_n)$, but the converse is not true. Nonetheless a uniform scaling by a constant α of the Hessian of an adjustment function A gives a more ambiguity averse preference if $\alpha > 1$. Therefore, to evaluate the average attitude of the DM toward ambiguity, we propose the widely used Frobenius norm which is:

$$\|\nabla^2 A(0_n)\|_F^2 = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \left| \frac{\partial^2 A}{\partial x_i \partial x_j}(0_n) \right|^2$$

and in the case of a symmetric matrix (Golub and Van Loan, 1996, formula 2.5.7):

$$\|\nabla^2 A(0_n)\|_F^2 = \|\Lambda\|_F^2 = \sum_{i=0}^{n-1} |\lambda_i|^2.$$

Setting $\Lambda' = \Lambda / \|\Lambda\|_F = \text{diag}(\lambda'_0, \dots, \lambda'_{n-1})$ with $\lambda'_i = -\lambda_i / \|\Lambda\|_F$, equation (11) rewrites as:

$$-\frac{1}{2} u'(w) \|\Lambda\|_F \langle \Lambda' \cdot \hat{\varphi}(h), \hat{\varphi}(h) \rangle \quad (12)$$

where the scalar product blends the variability of the utility profile of h and the DM specific levels of aversion to each sources of ambiguity. While the variability of h could easily be measured by the standard euclidian norm in \mathbb{R}^n , it is impossible to disentangle these two effects, indeed, they can compensate each other and one can only infer from the DM preferences the result of their product. Therefore we propose that, instead of considering the coordinates $\hat{\varphi}(h)$ of the projection of h in the rotated basis of the subspace of non crisp act, we use a “scaled” projection where the scaling along each source of ambiguity is proportional to the DM level of aversion to that source of ambiguity.

Formally define the linear operator from \mathbf{L}^2 to $\mathbf{NC} \subset \mathbf{L}^2$:

$$\mathbb{A}: h \mapsto \mathbb{A}h = \sum_{i=0}^{n-1} \sqrt{\lambda'_i} \mathbf{E}_{\mathbf{P}}(\hat{\zeta}_i \cdot h) \hat{\zeta}_i$$

We will call $\mathbb{A}h$ the *scaled non crisp projection* of h . Then the scalar product in equation (12) is:

$$\langle \Lambda' \cdot \hat{\varphi}(h), \hat{\varphi}(h) \rangle = \sum_{i=0}^{n-1} \lambda'_i (\hat{\varphi}_i(h))^2 = \sum_{i=0}^{n-1} (\sqrt{\lambda'_i} \hat{\varphi}_i(h))^2 = \sum_{i=0}^{n-1} (\sqrt{\lambda'_i} \mathbf{E}_{\mathbf{P}}(\hat{\zeta}_i \cdot h))^2$$

Using Parseval’s identity (Conway, 1990, Theorem 4.13(f)) we get:

$$\langle \Lambda' \cdot \hat{\varphi}(h), \hat{\varphi}(h) \rangle = \left\| \sum_{i=0}^{n-1} \sqrt{\lambda'_i} \mathbf{E}_{\mathbf{P}}(\hat{\zeta}_i \cdot h) \hat{\zeta}_i \right\|_{\mathbf{L}^2}^2 = \|\mathbb{A}h\|_{\mathbf{L}^2}^2$$

As $\mathbf{1}_{\mathbf{S}} \in \mathbf{C}$ implies $\mathbf{E}_{\mathbf{P}}(\mathbf{1}_{\mathbf{S}} \cdot \mathbb{A}h) = \mathbf{E}_{\mathbf{P}}(\mathbb{A}h) = 0$ we can conclude that:

$$\langle \Lambda' \cdot \hat{\varphi}(h), \hat{\varphi}(h) \rangle = \mathbf{var}_{\mathbf{P}}(\mathbb{A}h) \quad (13)$$

This equality explicits the relationship between variance and variation of the utility profile in the VEU model and allows to write the second-order Taylor-Young expansion of the certainty equivalent for an ambiguity averse DM as:

$$c(w+h) = w + \mathbf{E}_{\mathbf{P}}(h) - \frac{1}{2}\gamma(w) \mathbf{var}_{\mathbf{P}}(h) - \frac{1}{2}u'(w) \|\Lambda\|_{\mathbf{F}} \mathbf{var}_{\mathbf{P}}(\mathbb{A}h) + o(\|h\|_{\infty}^2) \quad (14)$$

Special cases. Two special cases where the separation of the DM attitude toward ambiguity and the characteristics of the act is achieved help illustrate this result. First, the case where a DM has equal aversion to each source of ambiguity, which implies that the matrix Λ' is the identity matrix, hence that the operator \mathbb{A} reduces to the projection on the subspace \mathbf{NC} of non crisp acts. Using the orthogonal decomposition $h = h_{\mathbf{C}} + h_{\mathbf{NC}}$ in $\mathbf{L}^2 = \mathbf{C} \oplus \mathbf{NC}$ with $h_{\mathbf{C}} \in \mathbf{C}$ and $h_{\mathbf{NC}} \in \mathbf{NC}$, we have that $\mathbf{var}_{\mathbf{P}}(\mathbb{A}h)$ is the variance of the purely non crisp component of the act: $\mathbf{var}_{\mathbf{P}}(h_{\mathbf{NC}})$ and the ambiguity aversion coefficient in equation (14) is $\frac{1}{2}u'(w) \|\Lambda\|_{\mathbf{F}} \mathbf{var}_{\mathbf{P}}(h_{\mathbf{NC}})$. Second, the case of an “ambiguity isotropic” act h for which there exists $\alpha \in \mathbb{R}_+^*$ such that $h_{\mathbf{NC}} = \sum_{i=0}^{n-1} \alpha \hat{\zeta}_i$ i.e. $\hat{\varphi}(h) = \alpha[1]$ (where $[1]$ is the vector made of 1s). Then we have $\mathbf{var}_{\mathbf{P}}(\mathbb{A}h) = \langle \Lambda' \hat{\varphi}(h), \hat{\varphi}(h) \rangle = \alpha^2 \mathbf{Tr} \Lambda'$ and as $n\alpha^2 = \mathbf{var}_{\mathbf{P}}(h_{\mathbf{NC}})$, the ambiguity aversion coefficient is $\frac{1}{2}u'(w) \frac{1}{n} \mathbf{Tr} \Lambda \mathbf{var}_{\mathbf{P}}(h_{\mathbf{NC}})$. In these two special cases the variance of the scaled non crisp projection of h is equal, up to a positive coefficient, to the variance of the non crisp component of h .

4.2 The Mean Variance Variability Criterion

From equation (14), the following mean variance preference generalised to take into account the variability of utility profiles is deduced: the DM ranks prospects $h \in \mathbf{L}^{\infty}$ according to the function

$$u_{\text{MNV}}(h) = \mathbf{E}_{\mathbf{P}}(h) - \frac{\gamma}{2} \mathbf{var}_{\mathbf{P}}(h) - \frac{\theta}{2} \mathbf{var}_{\mathbf{P}}(\mathbb{A}h) \quad (15)$$

where γ and θ are strictly positive coefficients measuring the respective aversion to risk and aversion to ambiguity.

Therefore a DM using the Mean Variance Variability preference measures the utility of a prospect by its expected utility reduced by a risk aversion factor proportional to its variance and by a risk ambiguity factor proportional to the variance of its scaled non crisp projection.

4.3 Comparison with Maccheroni, Marinacci, and Ruffino’s criterion

Maccheroni et al. (Forthcoming) use the smooth model of decision making under ambiguity of Klibanoff et al. (2005) to derive the Robust Mean-Variance Preference given by:

$$C(f) = \mathbf{E}_P(f) - \frac{\gamma}{2} \mathbf{var}_P(f) - \frac{\theta}{2} \sigma_{\mu}^2(\mathbf{E}(f))$$

Both their model and ours generalise the classical mean-variance preferences by subtracting a third term which accounts for the DM aversion to ambiguity and

which is proportional to a variance. In the Robust Mean-Variance Preference, the variance $\sigma_\mu^2(\mathbf{E}(f))$ is that of the expected utility $\mathbf{E}(f)$ (viewed as a random variable on the set of probabilities that have square integrable density with respect to the baseline probability P) calculated under the second order Borel probability μ . In the Mean Variance Variability criterion, the variance $\mathbf{var}_\mathbf{p}(\mathbb{A}h)$ is that of the scaled purely ambiguous part of f calculated under the endogenously determined baseline probability \mathbf{p} . Hence to apply the former model, the second order probability is needed while to apply the latter, the processes driving the independent sources of ambiguity have to be inferred.

The formal similarity between the two criteria explains why [Maccheroni et al.](#)'s results are easily transposed to our setting, nonetheless the properties of the VEU adjustment factors allow to further extend their analysis as will be seen in the next section.

5 Application to the choice of the Optimal Portfolio

In this application we suppose that the baseline probability \mathbf{p} has been inferred from the DM preferences over complementary acts and that all the random variables' moments considered are calculated under this baseline probability.

We consider a $k = 0$ risk-free asset whose return is r_0 and $k = 1, \dots, m$ risky (and possibly ambiguous) linearly independent assets whose vector of excess returns over r_0 , written $[\tilde{r}_k]$, is characterised by a vector of expected excess returns $[E_k] = [\mathbf{E}(\tilde{r}_k)]$ and a variance-covariance matrix $\Omega = [\sigma_{kl}]_{\substack{1 \leq k \leq m \\ 1 \leq l \leq m}}$. In our extended setting, sensibilities of these assets to ambiguity are measured by an (n, m) matrix Z (capital zeta) whose columns are the vectors $Z_k = \hat{\varphi}(\tilde{r}_k) = [\hat{\varphi}_i(\tilde{r}_k)]$ and whose elements are $Z_{ik} = \hat{\varphi}_i(\tilde{r}_k) = \mathbf{E}_\mathbf{p}(\hat{\zeta}_i \cdot \tilde{r}_k)$.

A portfolio is described by the vector $\psi \in \mathbb{R}^m$ of the proportions of the total wealth invested in each risky asset. The excess returns of the portfolios are given by $\tilde{r}_\psi = \psi^\top [\tilde{r}_k]$ and they are ranked by the DM using the Mean Variance Variability criterion: $u_{\text{MVV}}(\tilde{r}_\psi) = \mathbf{E}(\tilde{r}_\psi) - \frac{\gamma}{2} \mathbf{var}(\tilde{r}_\psi) - \frac{\theta}{2} \mathbf{var}(\mathbb{A}\tilde{r}_\psi)$. The expected excess return of the portfolio is $\mathbf{E}(\tilde{r}_\psi) = \psi^\top [E_k]$ and its variance is $\mathbf{var}(\tilde{r}_\psi) = \psi^\top \Omega \psi$. The coordinates in the subspace of non crisp acts of the return of the portfolio are given by $\hat{\varphi}_i(\tilde{r}_\psi) = \mathbf{E}_\mathbf{p}(\hat{\zeta}_i \cdot \tilde{r}_\psi) = \sum_{k=1}^m \psi_k \mathbf{E}_\mathbf{p}(\hat{\zeta}_i \cdot \tilde{r}_k)$ so that the vector of coordinates $\hat{\varphi}(\tilde{r}_\psi)$ is given by $Z\psi$. Therefore the ambiguity adjustment given by equation (13) is $\mathbf{var}(\mathbb{A}\tilde{r}_\psi) = \langle \Lambda' \cdot \hat{\varphi}(\tilde{r}_\psi), \hat{\varphi}(\tilde{r}_\psi) \rangle = (Z\psi)^\top \Lambda' Z\psi = \psi^\top Z^\top \Lambda' Z\psi$. Finally, the Mean Variance Variability criterion applied to the return of the portfolio can be written as follows:

$$u_{\text{MVV}}(\tilde{r}_\psi) = r_0 + \psi^\top [E_k] - \frac{\gamma}{2} \psi^\top \hat{\Omega} \psi \quad \text{where } \hat{\Omega} = \Omega + \frac{\theta}{\gamma} Z^\top \Lambda' Z \quad (16)$$

5.1 Analysis of the modified variance covariance matrix $\hat{\Omega}$.

Denoting by $\sqrt{\Lambda'} = \text{diag}(\sqrt{\lambda'_0}, \dots, \sqrt{\lambda'_{n-1}})$, we have $Z^\top \Lambda' Z = Z^\top (\sqrt{\Lambda'})^\top \sqrt{\Lambda'} Z = (\sqrt{\Lambda'} Z)^\top \sqrt{\Lambda'} Z$, hence $Z^\top \Lambda' Z$ is symmetric semi-definite positive. Recall that $Z = [\hat{\varphi}(\tilde{r}_1), \dots, \hat{\varphi}(\tilde{r}_m)]$, then the element (k, l) of $Z^\top \Lambda' Z$ is $\sum_{i=0}^{n-1} \lambda'_i \hat{\varphi}_i(\tilde{r}_k) \hat{\varphi}_i(\tilde{r}_l)$.

But

$$\begin{aligned}
\mathbf{cov}(\mathbb{A}\tilde{r}_k, \mathbb{A}\tilde{r}_l) &= \mathbf{E}(\mathbb{A}\tilde{r}_k \mathbb{A}\tilde{r}_l) \\
&= \mathbf{E} \left(\left(\sum_{i=0}^{n-1} \sqrt{\lambda'_i} \mathbf{E}_P(\hat{\zeta}_i \tilde{r}_k) \hat{\zeta}_i \right) \left(\sum_{j=0}^{n-1} \sqrt{\lambda'_j} \mathbf{E}_P(\hat{\zeta}_j \tilde{r}_l) \hat{\zeta}_j \right) \right) \\
&= \mathbf{E} \left(\sum_{i=0}^{n-1} \lambda'_i \mathbf{E}_P(\hat{\zeta}_i \tilde{r}_k) \mathbf{E}_P(\hat{\zeta}_i \tilde{r}_l) \hat{\zeta}_i^2 \right) = \sum_{i=0}^{n-1} \lambda'_i \hat{\varphi}_i(\tilde{r}_k) \hat{\varphi}_i(\tilde{r}_l)
\end{aligned}$$

Therefore $Z^\top \Lambda' Z$ is the variance covariance matrix of the scaled non crisp projections of the excess returns of the assets. The element (k, l) of the modified variance covariance matrix $\hat{\Omega}$ is:

$$\hat{\sigma}_{kl} = \mathbf{cov}(\tilde{r}_k, \tilde{r}_l) + \frac{\theta}{\gamma} \mathbf{cov}(\mathbb{A}\tilde{r}_k, \mathbb{A}\tilde{r}_l)$$

and especially, the diagonal terms are:

$$\hat{\sigma}_{kk} = \mathbf{var}(\tilde{r}_k) + \frac{\theta}{\gamma} \mathbf{var}(\mathbb{A}\tilde{r}_k)$$

hence we have $\hat{\sigma}_{kk} \geq \mathbf{var}(\tilde{r}_k)$.

The coefficients γ and θ being strictly positive and Ω being symmetric definite positive, the matrix $\hat{\Omega}$ is also symmetric definite positive, hence it is associated with a definite quadratic form. Therefore the modified variance covariance matrix can replace the variance covariance matrix in the purely risky applications to take into account ambiguity, and all the results concerning the efficient portfolio frontier or the CAPM can be rewritten with these modified values. Looking at the choice of the optimal portfolio, the first order condition for optimality when searching for the maximum of expression (16) with no constraint on ψ is:

$$\gamma \hat{\Omega} \psi = [E_k] \quad (17)$$

$\hat{\Omega}$ is definite positive hence invertible and the solution to equation (17) is given by:

$$\psi^* = \frac{1}{\gamma} \hat{\Omega}^{-1} [E_k] \quad (18)$$

5.2 The case of one ambiguous asset

When $m = 1$, *i.e.* there is one ambiguous asset and the risk-free asset, the solution (18) is:

$$\psi_1^* = \frac{E_1}{\gamma \hat{\sigma}_{11}} \quad (19)$$

where $\hat{\sigma}_{11} = \mathbf{var}(\tilde{r}_1) + \frac{\theta}{\gamma} \mathbf{var}(\mathbb{A}\tilde{r}_1)$. As expected, if \tilde{r}_1 is crisp, the formula reduces to the standard mean-variance solution (Markowitz, 1952). When the return of the asset is not crisp, it shows that an increase in the ambiguity of the asset or an increase in the aversion to ambiguity will lead to a higher demand for the risk-free asset.

5.3 The case of one risky and one ambiguous asset

In the case where asset 1 is risky but crisp and asset 2 is risky and non crisp, the modified variance covariance reduces to:

$$\widehat{\Omega} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \hat{\sigma}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{var}(\tilde{r}_1) & \mathbf{cov}(\tilde{r}_1, \tilde{r}_2) \\ \mathbf{cov}(\tilde{r}_1, \tilde{r}_2) & \mathbf{var}(\tilde{r}_2) + \frac{\theta}{\gamma} \mathbf{var}(\mathbb{A}\tilde{r}_2) \end{bmatrix}$$

and the solutions are:

$$\psi_1^* = \frac{1}{\gamma} \frac{E_1 \hat{\sigma}_{22} - E_2 \sigma_{12}}{\sigma_{11} \hat{\sigma}_{22} - \sigma_{12}^2} \quad \psi_2^* = \frac{1}{\gamma} \frac{E_2 \sigma_{11} - E_1 \sigma_{12}}{\sigma_{11} \hat{\sigma}_{22} - \sigma_{12}^2}$$

with the ratio of the weight of the risky asset over the weight of the ambiguous asset equal to:

$$\frac{\psi_1^*}{\psi_2^*} = \frac{E_1 \hat{\sigma}_{22} - E_2 \sigma_{12}}{E_2 \sigma_{11} - E_1 \sigma_{12}}$$

As announced in section 4.3, we can now prove that [Maccheroni et al. \(Forthcoming\)](#)'s asset allocation results hold in our axiomatically different setting. Namely define $(\alpha, \beta) = \operatorname{argmin} \|\tilde{r}_2 - (\alpha + \beta \tilde{r}_1)\|_{L^2}$, the coefficients of the linear regression of the excess return of the ambiguous asset 2 over the excess return of the risky asset 1 using the ordinary least square estimation:

$$\beta(\tilde{r}_1, \tilde{r}_2) = \frac{\mathbf{cov}(\tilde{r}_1, \tilde{r}_2)}{\mathbf{var}(\tilde{r}_1)} \quad \text{and} \quad \alpha(\tilde{r}_1, \tilde{r}_2) = E_2 - \beta(\tilde{r}_1, \tilde{r}_2) E_1$$

Then we have the following relationships:

Proposition 5.1. *Proportion of the ambiguous asset:*

$$\operatorname{sgn} \psi_2^* = \operatorname{sgn} \alpha(\tilde{r}_1, \tilde{r}_2)$$

Change in ambiguity aversion θ :

$$\begin{aligned} \operatorname{sgn} \frac{\partial}{\partial \theta} \left(\frac{\psi_1^*}{\psi_2^*} \right) &= \operatorname{sgn} \alpha(\tilde{r}_1, \tilde{r}_2) \\ \operatorname{sgn} \frac{\partial \psi_2^*}{\partial \theta} &= -\operatorname{sgn} \alpha(\tilde{r}_1, \tilde{r}_2), \quad \operatorname{sgn} \frac{\partial \psi_1^*}{\partial \theta} = \operatorname{sgn} \alpha(\tilde{r}_1, \tilde{r}_2) \beta(\tilde{r}_1, \tilde{r}_2) \end{aligned}$$

Change in risk aversion γ :

$$\operatorname{sgn} \frac{\partial}{\partial \gamma} \left(\frac{\psi_1^*}{\psi_2^*} \right) = \operatorname{sgn} \frac{\partial \psi_2^*}{\partial \gamma} = -\operatorname{sgn} \alpha(\tilde{r}_1, \tilde{r}_2)$$

This proves that the DM will hold a long position in the ambiguous asset if $\alpha(\tilde{r}_1, \tilde{r}_2)$ is positive and a short position otherwise. As this coefficient is positive when there is an expected added positive return in holding one unit of the ambiguous asset instead of $\beta(\tilde{r}_1, \tilde{r}_2)$ units of the risky asset, the investor chooses to trade off between a portfolio with a lower ambiguity and a potential larger return. The two sides of this trade-off are weighted by her aversion parameters as, in all cases, she will decrease in absolute terms the ambiguous asset holding when her ambiguity aversion θ increases. It is also noticeable that the variation

of the ratio of the two assets for a change in risk aversion and for a change in ambiguity aversion are of opposite signs, a relationship first obtained with the smooth model of decision making under ambiguity from simulations by [Klibanoff et al. \(2005, Example 2\)](#) and confirmed by [Maccheroni et al. \(Forthcoming\)](#). This paper proves that this inverse relationship also holds in the axiomatic setting of the Vector Expected Utility.

These results are a sound foundation for a discussion of the home-bias puzzle: replacing “risky asset” by “domestic asset” and “ambiguous asset” by “foreign asset” in the above paragraph gives an explanation through the investors aversion to ambiguity of the lack of diversification of international portfolios first exposed by [French and Poterba \(1991\)](#). Nonetheless it doesn’t seem necessary that the domestic asset should be purely risky. Therefore we propose in the next section an analysis of the case where the two assets are ambiguous, with the foreign asset either “objectively” more ambiguous than the domestic one (as measured by the \mathbb{R}^n norm of their vectors of non crisp coordinates) or perceived by the DM as more ambiguous (as measured by its correlation with sources of ambiguity associated with larger eigenvalues of the Hessian). To the best of our knowledge, this analysis has not yet been undertaken while the analytical tractability of our setting rendered it possible.

5.4 The case of two ambiguous assets

In the case where the two assets are not crisp, the modified variance covariance matrix is:

$$\hat{\Omega} = \begin{bmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{12} \\ \hat{\sigma}_{21} & \hat{\sigma}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{var}(\tilde{r}_1) + \frac{\theta}{\gamma} \mathbf{var}(\mathbb{A}\tilde{r}_1) & \mathbf{cov}(\tilde{r}_1, \tilde{r}_2) + \frac{\theta}{\gamma} \mathbf{cov}(\mathbb{A}\tilde{r}_1, \mathbb{A}\tilde{r}_2) \\ \mathbf{cov}(\tilde{r}_1, \tilde{r}_2) + \frac{\theta}{\gamma} \mathbf{cov}(\mathbb{A}\tilde{r}_1, \mathbb{A}\tilde{r}_2) & \mathbf{var}(\tilde{r}_2) + \frac{\theta}{\gamma} \mathbf{var}(\mathbb{A}\tilde{r}_2) \end{bmatrix}$$

and the solutions are:

$$\psi_1^* = \frac{1}{\gamma} \frac{E_1 \hat{\sigma}_{22} - E_2 \hat{\sigma}_{12}}{\hat{\sigma}_{11} \hat{\sigma}_{22} - \hat{\sigma}_{12}^2} \quad \psi_2^* = \frac{1}{\gamma} \frac{E_2 \hat{\sigma}_{11} - E_1 \hat{\sigma}_{12}}{\hat{\sigma}_{11} \hat{\sigma}_{22} - \hat{\sigma}_{12}^2} \quad (20)$$

We will denote by $A = E_1 \hat{\sigma}_{22} - E_2 \hat{\sigma}_{12}$, $B = E_2 \hat{\sigma}_{11} - E_1 \hat{\sigma}_{12}$ and $C = \hat{\sigma}_{11} \hat{\sigma}_{22} - \hat{\sigma}_{12}^2$ so that $\psi_1^* = \frac{1}{\gamma} \frac{A}{C}$ and $\psi_2^* = \frac{1}{\gamma} \frac{B}{C}$.

To analyze these results we need to decompose the excess returns into their respective crisp and ambiguous parts. Recall that in \mathbf{L}^2 , the subspaces \mathbf{C} of crisp acts and \mathbf{NC} of non crisp acts are orthogonal complement ([Siniscalchi, 2009](#), p. 844), then we can write $\tilde{r}_1 = \tilde{r}_{\mathbf{C}1} + \tilde{r}_{\mathbf{NC}1}$, $\tilde{r}_2 = \tilde{r}_{\mathbf{C}2} + \tilde{r}_{\mathbf{NC}2}$ with $\tilde{r}_{\mathbf{C}1}, \tilde{r}_{\mathbf{C}2} \in \mathbf{C}$, $\tilde{r}_{\mathbf{NC}1}, \tilde{r}_{\mathbf{NC}2} \in \mathbf{NC}$ and $\tilde{r}_{\mathbf{NC}1} = \langle Z_1, \hat{\zeta} \rangle$, $\tilde{r}_{\mathbf{NC}2} = \langle Z_2, \hat{\zeta} \rangle$. Given the orthogonality of \mathbf{C} and \mathbf{NC} , we have, for $i = \{1, 2\}$:

$$\begin{aligned} \mathbf{var}(\tilde{r}_i) &= \mathbf{var}(\tilde{r}_{\mathbf{C}i}) + \mathbf{var}(\tilde{r}_{\mathbf{NC}i}) \text{ with } \mathbf{var}(\tilde{r}_{\mathbf{NC}i}) = \|Z_i\|_{\mathbb{R}^n}^2 \\ \mathbf{cov}(\tilde{r}_1, \tilde{r}_2) &= \mathbf{cov}(\tilde{r}_{\mathbf{C}1}, \tilde{r}_{\mathbf{C}2}) + \mathbf{cov}(\tilde{r}_{\mathbf{NC}1}, \tilde{r}_{\mathbf{NC}2}) \text{ with } \mathbf{cov}(\tilde{r}_{\mathbf{NC}1}, \tilde{r}_{\mathbf{NC}2}) = \langle Z_1, Z_2 \rangle \\ \mathbf{var}(\mathbb{A}\tilde{r}_i) &= \left\| \sqrt{\Lambda'} Z_i \right\|_{\mathbb{R}^n}^2 \text{ and } \mathbf{cov}(\mathbb{A}\tilde{r}_1, \mathbb{A}\tilde{r}_2) = \langle \Lambda' Z_1, Z_2 \rangle \end{aligned}$$

Using some straightforward notations: $E_i = \mathbf{E}(\tilde{r}_i) = \mathbf{E}(\tilde{r}_{\mathbf{C}i})$, $\sigma_i^2 = \mathbf{var}(\tilde{r}_i)$, $\sigma_{\mathbf{C}i}^2 = \mathbf{var}(\tilde{r}_{\mathbf{C}i})$, $\sigma_{\mathbf{NC}i}^2 = \mathbf{var}(\tilde{r}_{\mathbf{NC}i})$, $\sigma_{\mathbb{A}i}^2 = \mathbf{var}(\mathbb{A}\tilde{r}_i)$ and $\mathbf{cov}(\tilde{r}_{\mathbf{C}1}, \tilde{r}_{\mathbf{C}2}) = \rho_{\mathbf{C}} \sigma_{\mathbf{C}1} \sigma_{\mathbf{C}2}$,

$\mathbf{cov}(\tilde{r}_{\text{NC1}}, \tilde{r}_{\text{NC2}}) = \rho_{\text{NC}} \sigma_{\text{NC1}} \sigma_{\text{NC2}}$, $\mathbf{cov}(\mathbb{A}\tilde{r}_1, \mathbb{A}\tilde{r}_2) = \rho_{\mathbb{A}} \sigma_{\mathbb{A1}} \sigma_{\mathbb{A2}}$, the numerators and the denominator of the solutions to the optimal portfolio problem can then be detailed:

Proposition 5.2. *In equation (20), the numerators are:*

$$\begin{aligned} A &= (E_1 \sigma_{\text{C2}}^2 - E_2 \rho_{\text{C}} \sigma_{\text{C1}} \sigma_{\text{C2}}) + (E_1 \sigma_{\text{NC2}}^2 - E_2 \rho_{\text{NC}} \sigma_{\text{NC1}} \sigma_{\text{NC2}}) + \frac{\theta}{\gamma} (E_1 \sigma_{\mathbb{A2}}^2 - E_2 \rho_{\mathbb{A}} \sigma_{\mathbb{A1}} \sigma_{\mathbb{A2}}) \\ B &= (E_2 \sigma_{\text{C1}}^2 - E_1 \rho_{\text{C}} \sigma_{\text{C1}} \sigma_{\text{C2}}) + (E_2 \sigma_{\text{NC1}}^2 - E_1 \rho_{\text{NC}} \sigma_{\text{NC1}} \sigma_{\text{NC2}}) + \frac{\theta}{\gamma} (E_2 \sigma_{\mathbb{A1}}^2 - E_1 \rho_{\mathbb{A}} \sigma_{\mathbb{A1}} \sigma_{\mathbb{A2}}) \end{aligned}$$

and the denominator is:

$$\begin{aligned} C &= \sigma_{\text{C1}}^2 \sigma_{\text{C2}}^2 (1 - \rho_{\text{C}}^2) + \sigma_{\text{NC1}}^2 \sigma_{\text{NC2}}^2 (1 - \rho_{\text{NC}}^2) + \frac{\theta^2}{\gamma^2} \sigma_{\mathbb{A1}}^2 \sigma_{\mathbb{A2}}^2 (1 - \rho_{\mathbb{A}}^2) \\ &\quad + (\sigma_{\text{C1}} \sigma_{\text{NC2}} - \sigma_{\text{C2}} \sigma_{\text{NC1}})^2 + 2 \sigma_{\text{C1}} \sigma_{\text{C2}} \sigma_{\text{NC1}} \sigma_{\text{NC2}} (1 - \rho_{\text{C}} \rho_{\text{NC}}) \\ &\quad + \frac{\theta}{\gamma} [(\sigma_{\text{C1}} \sigma_{\mathbb{A2}} - \sigma_{\text{C2}} \sigma_{\mathbb{A1}})^2 + 2 \sigma_{\text{C1}} \sigma_{\text{C2}} \sigma_{\mathbb{A1}} \sigma_{\mathbb{A2}} (1 - \rho_{\text{C}} \rho_{\mathbb{A}})] \\ &\quad + \frac{\theta}{\gamma} [(\sigma_{\mathbb{A1}} \sigma_{\text{NC2}} - \sigma_{\mathbb{A2}} \sigma_{\text{NC1}})^2 + 2 \sigma_{\text{NC1}} \sigma_{\text{NC2}} \sigma_{\mathbb{A1}} \sigma_{\mathbb{A2}} (1 - \rho_{\text{NC}} \rho_{\mathbb{A}})] \end{aligned}$$

It follows directly that $C \geq 0$ hence the choice of a long or a short position in the two assets is determined by the signs of the numerators A and B . These formulae also highlight the pivotal role of the correlations (of the crisp parts and of the ambiguous parts) of the excess returns in determining the composition of the optimal portfolio.

We will now study the case where each asset is correlated to a distinct source of ambiguity, therefore their non crisp components are given by $\tilde{r}_{\text{NC1}} = Z_{11} \hat{\zeta}_1$ and $\tilde{r}_{\text{NC2}} = Z_{02} \hat{\zeta}_0$. As the sources are independent this implies $\rho_{\text{NC}} = \rho_{\mathbb{A}} = 0$ that is there is no possibility to hedge away some or all of the ambiguity by holding the two assets. The DM has again to trade off expected return, risk and ambiguity following the Mean Variance Variability criterion which represents her utility. To further focus on the effects of ambiguity, we will consider two assets whose excess returns have similar crisp parts with the same expected return $E_1 = E_2 = E$ and the same variance $\sigma_{\text{C1}}^2 = \sigma_{\text{C2}}^2 = \sigma_{\text{C}}^2$. Recall that the DM is more averse to the source of ambiguity $\hat{\zeta}_0$ than to the source $\hat{\zeta}_1$ ($\lambda'_0 > \lambda'_1$), hence one may consider in an home-bias puzzle setting that asset 1 is the domestic one and asset 2 the foreign one.

From proposition 5.2 we get:

$$\begin{aligned} A &= E \left[\sigma_{\text{C}}^2 (1 - \rho_{\text{C}}) + Z_{02}^2 \left(1 + \frac{\theta}{\gamma} \lambda'_0 \right) \right] \\ B &= E \left[\sigma_{\text{C}}^2 (1 - \rho_{\text{C}}) + Z_{11}^2 \left(1 + \frac{\theta}{\gamma} \lambda'_1 \right) \right] \end{aligned}$$

Then as $\psi_1^* / \psi_2^* = A/B$:

$$\begin{aligned} \text{sgn} \frac{\partial}{\partial \theta} \left(\frac{\psi_1^*}{\psi_2^*} \right) &= \text{sgn} [\sigma_{\text{C}}^2 (1 - \rho_{\text{C}}) (\lambda'_0 Z_{11}^2 - \lambda'_1 Z_{02}^2) + Z_{11}^2 Z_{02}^2 (\lambda'_0 - \lambda'_1)] \\ \text{sgn} \frac{\partial}{\partial \gamma} \left(\frac{\psi_1^*}{\psi_2^*} \right) &= \text{sgn} [\sigma_{\text{C}}^2 (1 - \rho_{\text{C}}) (\lambda'_1 Z_{02}^2 - \lambda'_0 Z_{11}^2) + Z_{11}^2 Z_{02}^2 (\lambda'_1 - \lambda'_0)] \end{aligned}$$

First, the inverse relationship between the sensitivity of the composition of the portfolio to ambiguity aversion and to risk aversion holds with two ambiguous

assets correlated to distinct sources of ambiguity. Second, we have that $\psi_1^* \geq \psi_2^*$ if and only if $Z_{02}^2(1 + \frac{\theta}{\gamma}\lambda'_0) \geq Z_{11}^2(1 + \frac{\theta}{\gamma}\lambda'_1)$, hence it is possible that asset 2 could get a larger share of the portfolio than asset 1 if the ratio of their exposures to ambiguity Z_{02}/Z_{11} was to compensate for the DM higher aversion to the source $\hat{\zeta}_0$. While this case is possible, it would not illustrate properly a home-bias application as it would mean comparing a high beta domestic stock to a low beta foreign stock. Therefore, we can sharpen the above results by assuming that $Z_{11}^2 = Z_{02}^2 = \sigma_{\text{NC}}^2$ which help focus the analysis on the DM different attitude toward the two sources of ambiguity. With this added assumption we have a larger holding of the domestic asset 1 while both assets have the same expected return E and the same variance $\sigma_C^2 + \sigma_{\text{NC}}^2$ and we have

$$\text{sgn} \frac{\partial}{\partial \theta} \left(\frac{\psi_1^*}{\psi_2^*} \right) = \text{sgn} [(\lambda'_0 - \lambda'_1)(\sigma_C^2(1 - \rho_C) + \sigma_{\text{NC}}^2)] = \text{sgn}(\lambda'_0 - \lambda'_1) > 0$$

hence an increase in ambiguity aversion will decrease the holding of the foreign asset 2. Therefore the salient facts of the home-bias puzzle are retrieved in our setting with two ambiguous assets, one perceived as more ambiguous than the other by the DM, instead of one risky and one ambiguous asset.

6 Conclusion

In this paper we have aimed to corroborate the proposition that ambiguity matters for modern portfolio theory inasmuch as it can help understand some of the otherwise puzzling features of empirical data. To this end, we have studied compositions of optimal portfolios when asset returns are ambiguous using a generalised mean-variance preference that we have founded on an approximation “in the small” of Siniscalchi (2009)’s Vector Expected Utility’s certainty equivalent. The VEU model has been chosen for its axiomatic foundations which are built upon the seminal work of Gilboa and Schmeidler (1989), extended by Ghirardato, Maccheroni, and Marinacci (2004) and Maccheroni, Marinacci, and Rustichini (2006) among others, papers with numerous applications to finance. Its axiomatic foundations also have a very appealing behavioural interpretation in asset markets: indeed, in this context complementary acts have a straightforward meaning (detailed in the introduction) and the cognitive assumption which underpins the central “Complementary Independence” axiom — complementary acts have the same variability — is easily subscribed to. Moreover the model permits an identification of a baseline prior from the DM preferences and accounts for eventwise complementarity which may allow for “ambiguity diversification”.

The calculus of the second-order Taylor-Young expansion of the VEU certainty equivalent has lead to a generalisation of the mean-variance preference through a new term which measures the aversion to ambiguity of the DM and complements the classical aversion to variance. Thanks to its analytical tractability, this *Mean Variance Variability* criterion has allowed to generalise Markowitz (1952)’ results: we have shown that aversion to ambiguity can be accounted for by a variance-covariance matrix modified by the second order moments of the *scaled non crisp projections* of the asset returns. Our findings have not only confirmed the robustness to axiomatic specifications of some of the literature’s results on

asset prices under ambiguity obtained with other models, but also have extended the existing analysis to multiple ambiguous assets, hence have allowed to propose a novel discussion of the home-bias puzzle modeled with two ambiguous assets. Among other results, we have shown that the inverse relationship between the sensitivities of the portfolio composition to aversion to risk and aversion to ambiguity still hold with two ambiguous assets and that the holding of the most ambiguous asset will decline with heightened aversion to ambiguity.

Finally this Mean Variance Variability criterion opens the way for further research on equilibrium asset prices, possibly in a continuous time setting, and to some empirical studies and calibrations which are facilitated by its simple formulation. Especially, a study of the drivers of assets returns and risks could help define the main independent sources of ambiguity which are central to the workings of VEU model.

Appendix A Proofs

A.1 Quadratic approximation

For the calculus of the differential, we introduce 2 additional functions:

$$\begin{aligned} \Pi: \mathbb{L}^\infty &\rightarrow \mathbb{R}^n \times \mathbb{R} & W: \mathbb{R}^n \times \mathbb{R} &\rightarrow \mathbb{R} \\ a &\mapsto \begin{pmatrix} \varphi(a) \\ \mathbf{E}_p(a) \end{pmatrix} & (\varphi, \lambda) &\mapsto \lambda + A(\varphi) \end{aligned}$$

With these additional notations, we can rewrite equation (3) as

$$V(f) = W \circ \Pi \circ U(f)$$

and the certainty equivalent as

$$c(f) = u^{-1} \circ W \circ \Pi \circ U(f).$$

Note that, in this section, to keep the notations as light as possible, $\langle a, b \rangle$ will be the scalar product of \mathbb{R}^n when the functions A and φ are involved and the scalar product of \mathbb{R}^{n+1} when the functions W and Π are involved.

Proof of Proposition 3.1 Take $\varepsilon \in \mathbb{R}_+^*$. As u is differentiable over \mathbb{R} , at any point x there exists $\alpha \in \mathbb{R}_+^*$ such that for all $y \in \mathbb{R}$, $|y| \leq \alpha$ implies $|u(x+y) - u(x) - y.u'(x)| \leq \varepsilon|y|$. Hence for a given $a \in \mathbb{L}^\infty$ and for any $h \in \mathbb{L}^\infty$ and $s \in S$ such that $|h(s)| \leq \alpha$ we have $|u(a(s) + h(s)) - u(a(s)) - h(s).u'(a(s))| \leq \varepsilon|h(s)|$.

Now choose h such that $\|h\|_\infty \leq \alpha$, and set $E_h := \{s \mid |h(s)| > \|h\|_\infty\}$ then by definition $\mathbf{p}(E_h) = 0$. For all $s \in S \setminus E_h$ we have $|h(s)| \leq \|h\|_\infty \leq \alpha$ and the previous inequality holds: $|u(a(s) + h(s)) - u(a(s)) - h(s).u'(a(s))| \leq \varepsilon|h(s)| \leq \varepsilon\|h\|_\infty$. As the lhs is bounded \mathbf{p} -almost everywhere by $\varepsilon\|h\|_\infty$ so is its essential supremum and we get the desired result:

$$\|U(a+h) - U(a) - h.U'(a)\|_\infty \leq \varepsilon\|h\|_\infty$$

which is U is differentiable and $d_a U(h) = hU'(a)$.

Using the same reasoning replacing u by u' which is also differentiable over \mathbb{R} , we get that for all $\varepsilon \in \mathbb{R}_+^*$ there exists $\alpha \in \mathbb{R}_+^*$ such that for all $k \in \mathbb{L}^\infty$ with $\|k\|_\infty \leq \alpha$ we have

$$\|U'(a+k) - U'(a) - k.U''(a)\|_\infty \leq \varepsilon \|k\|_\infty.$$

Then

$$\begin{aligned} \|dU(a+k) - dU(a) - kU''(a)\|_{\mathcal{L}(\mathbb{L}^\infty, \mathbb{L}^\infty)} &= \sup_{\|h\|_\infty \leq 1} \|dU(a+k)(h) - dU(a)(h) - hU''(a)\|_\infty \\ &= \sup_{\|h\|_\infty \leq 1} \|hU'(a+k) - hU'(a) - hU''(a)\|_\infty \end{aligned}$$

using Hölder's inequality⁸:

$$\begin{aligned} &\leq \sup_{\|h\|_\infty \leq 1} \|h\|_\infty \|U'(a+k) - U'(a) - U''(a)h\|_\infty \\ &\leq \|U'(a+k) - U'(a) - U''(a)h\|_\infty \\ &\leq \varepsilon \|k\|_\infty \end{aligned}$$

which is the result : U is twice differentiable and $d_a^2 U(k, h) = kU''(a)h$. \square

Proof of Proposition 3.2 Using the chain rule for differentials we get:

$$d_w c(h) = d_w(u^{-1} \circ W \circ \Pi \circ U)(h) = d_{W \circ \Pi \circ U(w)} u^{-1} \circ d_{\Pi \circ U(w)} W \circ d_{U(w)} \Pi \circ d_w U(h)$$

The differential of U is given by Proposition 3.1, the differential of Π , a linear mapping, is itself, the differential of W is given by its gradient thanks to Assumption 2 and the differential of u^{-1} is given by the following lemma:

Lemma A.1. *As u is continuously differentiable and strictly increasing, u^{-1} is \mathbb{L}^∞ -differentiable and for all $a, h \in \mathbb{L}^\infty$: $d_a u^{-1}(h) = h/u'(u^{-1}(a))$.*

Proof. As u is strictly monotone and continuous, it is an homomorphism, and being strictly monotone and differentiable, it is a diffeomorphism. Then its differential is an isomorphism whose inverse is $(d_b u)^{-1} = d_{u(b)}(u^{-1})$. Setting $b = u^{-1}(a)$ we have that $d_a(u^{-1}) \circ d_{u^{-1}(a)} u = Id$, that is for all $h \in \mathbb{R}$, $d_a(u^{-1}) \circ (h.u'(u^{-1}(a))) = h$. By the linearity of the differential we obtain

$$u'(u^{-1}(a)).d_a(u^{-1})(h) = h$$

which gives the result. \square

Assembling all the pieces, we get:

$$\begin{aligned} d_w U(h) &= hU'(w) \\ d_{U(w)} \Pi(hU'(w)) &= \Pi(hU'(w)) \\ d_{\Pi \circ U(w)} W(\Pi(hU'(w))) &= \langle \nabla W(\Pi(U(w))), \Pi(hU'(w)) \rangle \\ d_{W \circ \Pi \circ U(w)} u^{-1}(\langle \nabla W(\Pi(U(w))), \Pi(hU'(w)) \rangle) &= \frac{\langle \nabla W(\Pi(U(w))), \Pi(hU'(w)) \rangle}{u'(u^{-1}(W \circ \Pi \circ U(w)))} \end{aligned}$$

⁸See, for example, [Aliprantis and Border, 2006](#), Prop. 13.2.

As $u^{-1}(W \circ \Pi \circ U(w)) = c(w)$, this is:

$$d_w c(h) = \frac{\langle \nabla W(\Pi(U(w))), \Pi(hU'(w)) \rangle}{u'(c(w))} \quad (21)$$

From the definition of W we have $\nabla W = (\nabla A, 1)$. Using the notations $\hat{w} := \varphi(U(w))$ and $\dot{w} := U'(w)/u'(c(w))$, the preceding formula is:

$$\begin{aligned} d_w c(h) &= \frac{\mathbf{E}_{\mathbf{P}}(hU'(w)) + \langle \nabla A(\hat{w}), \varphi(hU'(w)) \rangle}{u'(c(w))} \\ &= \mathbf{E}_{\mathbf{P}} \left(h \frac{U'(w)}{u'(c(w))} \right) + \left\langle \nabla A(\hat{w}), \varphi \left(h \frac{U'(w)}{u'(c(w))} \right) \right\rangle \\ &= \mathbf{E}_{\mathbf{P}}(h\dot{w}) + \langle \nabla A(\hat{w}), \varphi(h\dot{w}) \rangle \quad \square \end{aligned}$$

Proof of Proposition 3.3 To obtain the second order differential, we need to differentiate the first-order differential function $dc: \mathbf{L}^\infty \rightarrow \mathcal{L}(\mathbf{L}^\infty, \mathbb{R})$ such that $dc(w) = d_w c$.

We introduce some auxiliary functions :

- $\alpha: \mathbf{L}^\infty \rightarrow \mathcal{L}(\mathbf{L}^\infty, \mathbb{R}^{n+1})$ such that $\alpha(w): \mathbf{L}^\infty \rightarrow \mathbb{R}^{n+1}$ is defined by

$$\alpha(w)(h) = \Pi(hU'(w)),$$

- $\beta: \mathbf{L}^\infty \rightarrow \mathbb{R} \times \mathbb{R}^{n+1} \times \mathcal{L}(\mathbf{L}^\infty, \mathbb{R}^{n+1})$ such that

$$\beta(w) = (u'(c(w)), \nabla W(\Pi(U(w))), \alpha(w)),$$

- $\gamma: \mathbb{R} \times \mathbb{R}^{n+1} \times \mathcal{L}(\mathbf{L}^\infty, \mathbb{R}^{n+1}) \rightarrow \mathcal{L}(\mathbf{L}^\infty, \mathbb{R})$ such that

$$\gamma(t, a, \varphi) = \frac{\langle a, \varphi \rangle}{t}.$$

Then, starting from equation (21), we have $dc = \gamma \circ \beta$ and $d_w^2 c = d_w(\gamma \circ \beta) = d_{\beta(w)} \gamma \circ d_w \beta$.

Step 1: $d_w \alpha$. Take $\varepsilon \in \mathbb{R}_+^*$. As U' is differentiable there exists $\alpha \in \mathbb{R}_+^*$ such that for all $k \in \mathbf{L}^\infty$ with $\|k\|_\infty \leq \alpha$ we have

$$\|U'(w+k) - U(w) - kU''(w)\|_\infty \leq \varepsilon \|k\|_\infty.$$

Then for any $h \in \mathbf{L}^\infty$, using Hölder's inequality:

$$\begin{aligned} \|h \cdot (U'(w+k) - U(w) - kU''(w))\|_\infty &\leq \|h\|_\infty \|U'(w+k) - U(w) - kU''(w)\|_\infty \\ &\leq \varepsilon \|h\|_\infty \|k\|_\infty \end{aligned}$$

As all norms are equivalent on \mathbb{R}^{n+1} , we can choose w.l.o.g. $\|\Pi(a)\|_{\mathbb{R}^{n+1}} = \sup_{0 \leq i \leq n} |\Pi_i(a)|$ for any $a \in \mathbf{L}^\infty$. Then

$$\|\Pi(a)\|_{\mathbb{R}^{n+1}} = \sup_{0 \leq i \leq n} |\mathbf{E}_{\mathbf{P}}[\zeta_i \cdot a]| \leq \sup_{0 \leq i \leq n} \|\zeta_i \cdot a\|_{\mathbf{L}^1}$$

if we set $\zeta_n = \mathbf{1}_S$. Using again Hölder's inequality we get $\|\zeta_i \cdot a\|_{L^1} \leq \|\zeta_i\|_{L^1} \|a\|_\infty$. Then a corollary of the same inequality (Aliprantis and Border, 2006, Corollary 13.3) states that $L^2(\mathbf{p}) \subset L^1(\mathbf{p})$ and, when $\mathbf{p}(\Omega) = 1$, $\|\zeta_i\|_{L^1} \leq \|\zeta_i\|_{L^2}$. Now recall that the $\{\zeta_i\}_{0 \leq i < n}$ have been obtained through the Gram-Schmidt orthogonalization process in L^2 (Siniscalchi, 2009, p. 843), hence for all $0 \leq i < n$, $\|\zeta_i\|_{L^2} = 1$, to conclude that $\sup_{0 \leq i \leq n} \|\zeta_i \cdot a\|_{L^1} \leq \|a\|_\infty$ and finally $\|\Pi(a)\|_{\mathbb{R}^{n+1}} \leq \|a\|_\infty$.

Combining the previous results we have:

$$\begin{aligned} \|\alpha(w+k)(h) - \alpha(w)(h) - \Pi(hkU''(w))\|_{\mathbb{R}^{n+1}} &= \|\Pi(hU'(w+k)) - \Pi(hU(w)) - \Pi(hkU''(w))\|_{\mathbb{R}^{n+1}} \\ &\leq \|hU'(w+k) - hU(w) - hkU''(w)\|_\infty \\ &\leq \varepsilon \|h\|_\infty \|k\|_\infty \end{aligned}$$

Then

$$\begin{aligned} \|\alpha(w+k) - \alpha(w) - \Pi(\cdot kU''(w))\|_{\mathcal{L}(L^\infty, \mathbb{R}^{n+1})} &= \sup_{\|h\|_\infty \leq 1} \|\alpha(w+k)(h) - \alpha(w)(h) - \Pi(hkU''(w))\|_{\mathbb{R}^{n+1}} \leq \varepsilon \|k\|_\infty \end{aligned}$$

which proves that the differential of α at point w in direction k is the function $d_w\alpha(k): L^\infty \rightarrow \mathcal{L}(L^\infty, \mathbb{R}^{n+1})$ such that $d_w\alpha(k, h) = \Pi(hkU''(w))$.

Step 2: $d_w\beta$. Using the rules of differentiation for functions with values in a product space we have :

$$d_w\beta = (d_w(u' \circ c), d_w(\nabla W \circ \Pi \circ U), d_w\alpha)$$

Where we have for all $k \in L^\infty$:

$$\begin{aligned} d_w(u' \circ c)(k) &= d_{c(w)}u' \circ d_w c(k) = u''(c(w)) \cdot d_w c(k) \\ &= \frac{u''(c(w))}{u'(c(w))} \langle \nabla W(\Pi(U(w))), \Pi(kU'(w)) \rangle \end{aligned}$$

$$\begin{aligned} d_w(\nabla W \circ \Pi \circ U)(k) &= d_{\Pi \circ U(w)} \nabla W \circ d_{U(w)} \Pi \circ d_w U(k) \\ &= d_{\Pi \circ U(w)} \nabla W(\Pi(kU'(w))) \\ &= \nabla^2 W(\Pi(U(w))) \cdot \Pi(kU'(w)) \end{aligned}$$

and $d_w\alpha$ has been obtained at step 1.

Step 3: $d_w\gamma$. We now need the differential of a function defined on a product space. Let (s, b, ψ) be a point in $\mathbb{R} \times \mathbb{R}^{n+1} \times \mathcal{L}(L^\infty, \mathbb{R}^{n+1})$. Then for any direction (t, a, φ) in the same product space define the functions :

$$\begin{aligned} \gamma_1: \mathbb{R} &\rightarrow \mathcal{L}(L^\infty, \mathbb{R}) \text{ such that } \gamma_1(t) = \gamma(t, b, \psi) = \langle b, \psi \rangle / t \\ \gamma_2: \mathbb{R}^{n+1} &\rightarrow \mathcal{L}(L^\infty, \mathbb{R}) \text{ such that } \gamma_2(a) = \gamma(s, a, \psi) = \langle a, \psi \rangle / s \\ \gamma_3: \mathcal{L}(L^\infty, \mathbb{R}^{n+1}) &\rightarrow \mathcal{L}(L^\infty, \mathbb{R}) \text{ such that } \gamma_3(\varphi) = \gamma(s, b, \varphi) = \langle b, \varphi \rangle / s \end{aligned}$$

so that we have

$$d_{(s,b,\psi)}\gamma(t,a,\varphi) = d_s\gamma_1(t) + d_b\gamma_2(a) + d_\psi\gamma_3(\varphi).$$

Now we compute the three partial differentials. The first one is obtained directly by the derivative: $d_s\gamma_1(t) = -\frac{\langle b, \psi \rangle}{s^2}t$. The second and third ones are direct consequences of the linearity of the scalar product, indeed we have $\gamma_2(b+a) = \gamma_2(b) + \gamma_2(a)$ hence $d_b\gamma_2(a) = \gamma_2(a) = \langle a, \psi \rangle / s$ and $\gamma_3(\psi + \varphi) = \gamma_3(\psi) + \gamma_3(\varphi)$ hence $d_\psi\gamma_3(\varphi) = \gamma_3(\varphi) = \langle b, \varphi \rangle / s$. To summarize:

$$d_{(s,b,\psi)}\gamma(t,a,\varphi) = -\frac{\langle b, \psi \rangle}{s^2}t + \frac{\langle a, \psi \rangle}{s} + \frac{\langle b, \varphi \rangle}{s}$$

Step 4: d_w^2c . Going back to second order differential we have:

$$\begin{aligned} d_w^2c(k) &= d_{\beta(w)}\gamma \circ d_w\beta(k) \\ &= d_{(u'(c(w)), \nabla W(\Pi(U(w))), \alpha(w))}\gamma \circ d_w\beta(k) \\ &= d_{(u'(c(w)), \nabla W(\Pi(U(w))), \alpha(w))}\gamma(d_w(u' \circ c)(k), d_w(\nabla W \circ \Pi \circ U)(k), d_w\alpha(k)) \end{aligned}$$

using Step 3:

$$\begin{aligned} d_w^2c(k) &= -\frac{\langle \nabla W(\Pi(U(w))), \alpha(w) \rangle}{(u'(c(w)))^2} d_w(u' \circ c)(k) \\ &\quad + \frac{\langle d_w(\nabla W \circ \Pi \circ U)(k), \alpha(w) \rangle}{u'(c(w))} \\ &\quad + \frac{\langle \nabla W(\Pi(U(w))), d_w\alpha(k) \rangle}{u'(c(w))} \end{aligned}$$

using Step 2:

$$\begin{aligned} d_w^2c(k) &= -\frac{u''(c(w))}{(u'(c(w)))^3} \langle \nabla W(\Pi(U(w))), \Pi(kU'(w)) \rangle \langle \nabla W(\Pi(U(w))), \alpha(w) \rangle \\ &\quad + \frac{1}{u'(c(w))} \langle \nabla^2 W(\Pi(U(w))).\Pi(kU'(w)), \alpha(w) \rangle \\ &\quad + \frac{1}{u'(c(w))} \langle \nabla W(\Pi(U(w))), d_w\alpha(k) \rangle \end{aligned}$$

Taking the value at h and using Step 1:

$$\begin{aligned} d_w^2c(k, h) &= -\frac{u''(c(w))}{(u'(c(w)))^3} \langle \nabla W(\Pi(U(w))), \Pi(kU'(w)) \rangle \langle \nabla W(\Pi(U(w))), \Pi(hU'(w)) \rangle \\ &\quad + \frac{1}{u'(c(w))} \langle \nabla^2 W(\Pi(U(w))).\Pi(kU'(w)), \Pi(hU'(w)) \rangle \\ &\quad + \frac{1}{u'(c(w))} \langle \nabla W(\Pi(U(w))), \Pi(hkU''(w)) \rangle \end{aligned}$$

Finally, when $k = h$ as needed in the Taylor-Young formula:

$$\begin{aligned} d_w^2 c(h)^2 &= -\frac{u''(c(w))}{(u'(c(w)))^3} (\langle \nabla W(\Pi(U(w))), \Pi(hU'(w)) \rangle)^2 \\ &\quad + \frac{1}{u'(c(w))} \langle \nabla^2 W(\Pi(U(w))).\Pi(hU'(w)), \Pi(hU'(w)) \rangle \\ &\quad + \frac{1}{u'(c(w))} \langle \nabla W(\Pi(U(w))), \Pi(h^2 U''(w)) \rangle \end{aligned}$$

As the Hessian matrix of W is the Hessian matrix of A bordered with zeros, we can express this result with the functions A and φ :

$$\begin{aligned} d_w^2 c(h)^2 &= -\frac{u''(c(w))}{(u'(c(w)))^3} (\mathbf{E}_{\mathbf{P}}(hU'(w)) + \langle \nabla A(\varphi(U(w))), \varphi(hU'(w)) \rangle)^2 \\ &\quad + \frac{1}{u'(c(w))} \langle \nabla^2 A(\varphi(U(w))).\varphi(hU'(w)), \varphi(hU'(w)) \rangle \\ &\quad + \frac{1}{u'(c(w))} (\mathbf{E}_{\mathbf{P}}(h^2 U''(w)) + \langle \nabla A(\varphi(U(w))), \varphi(h^2 U''(w)) \rangle) \end{aligned}$$

With $\hat{w} := \varphi(U(w))$ and distributing the denominators:

$$\begin{aligned} d_w^2 c(h)^2 &= -\frac{u''(c(w))}{u'(c(w))} \left(\mathbf{E}_{\mathbf{P}} \left(h \frac{U'(w)}{u'(c(w))} \right) + \left\langle \nabla A(\hat{w}), \varphi \left(h \frac{U'(w)}{u'(c(w))} \right) \right\rangle \right)^2 \\ &\quad + u'(c(w)) \left\langle \nabla^2 A(\hat{w}).\varphi \left(h \frac{U'(w)}{u'(c(w))} \right), \varphi \left(h \frac{U'(w)}{u'(c(w))} \right) \right\rangle \\ &\quad + \frac{u''(c(w))}{u'(c(w))} \left(\mathbf{E}_{\mathbf{P}} \left(h^2 \frac{U''(w)}{u''(c(w))} \right) + \left\langle \nabla A(\hat{w}), \varphi \left(h^2 \frac{U''(w)}{u''(c(w))} \right) \right\rangle \right) \end{aligned}$$

With the notations $\gamma(c(w)) := -u''(c(w))/u'(c(w))$, $\dot{w} := U'(w)/u'(c(w))$ and $\ddot{w} := U''(w)/u''(c(w))$ and developing the square:

$$\begin{aligned} d_w^2 c(h)^2 &= \gamma(c(w)) \left((\mathbf{E}_{\mathbf{P}}(h\dot{w}))^2 + 2 \mathbf{E}_{\mathbf{P}}(h\dot{w}) \langle \nabla A(\hat{w}), \varphi(h\dot{w}) \rangle + \langle \nabla A(\hat{w}), \varphi(h\dot{w}) \rangle^2 \right) \\ &\quad + u'(c(w)) \langle \nabla^2 A(\hat{w}).\varphi(h\dot{w}), \varphi(h\dot{w}) \rangle \\ &\quad - \gamma(c(w)) (\mathbf{E}_{\mathbf{P}}(h^2 \ddot{w}) + \langle \nabla A(\hat{w}), \varphi(h^2 \ddot{w}) \rangle) \end{aligned}$$

Finally rearranging the terms gives the result:

$$\begin{aligned} d_w^2 c(h)^2 &= -\gamma(c(w)) \left(\mathbf{E}_{\mathbf{P}}(h^2 \ddot{w}) - (\mathbf{E}_{\mathbf{P}}(h\dot{w}))^2 \right) \\ &\quad - \gamma(c(w)) \left(\langle \nabla A(\hat{w}), \varphi(h^2 \ddot{w}) \rangle - \langle \nabla A(\hat{w}), \varphi(h\dot{w}) \rangle^2 \right) \\ &\quad + 2\gamma(c(w)) \mathbf{E}_{\mathbf{P}}(h\dot{w}) \langle \nabla A(\hat{w}), \varphi(h\dot{w}) \rangle \\ &\quad + u'(c(w)) \langle \nabla^2 A(\hat{w}).\varphi(h\dot{w}), \varphi(h\dot{w}) \rangle \quad \square \end{aligned}$$

A.2 Mean Variance Variability Preference

Proof of Proposition 4.1 Proposition 4 in [Siniscalchi \(2009\)](#) proves that \succsim_1 is more ambiguity averse than \succsim_2 if and only if for all $d \in \mathbb{R}^n$, $A^1(d) \leq A^2(d)$.

Suppose that \succsim_1 is more ambiguity averse than \succsim_2 and that $|\lambda_i^1| < |\lambda_i^2|$ that is there exist $\eta > 0$ such that $\lambda_i^1 = \lambda_i^2 + 2\eta$ (all eigenvalues are negative). Take an act $h = \alpha \hat{e}_i$ with $\alpha \in \mathbb{R}^*$, that is $\hat{\varphi}(h) = \alpha \hat{e}_i$. Then with our assumptions, the second order approximations are: $A^1(\hat{\varphi}(h)) = \frac{1}{2} \lambda_i^1 \alpha^2 + o(|\alpha|^2)$ and $A^2(\hat{\varphi}(h)) = \frac{1}{2} \lambda_i^2 \alpha^2 + o(|\alpha|^2)$ and we have

$$A^2(\hat{\varphi}(h)) - A^1(\hat{\varphi}(h)) + \eta \alpha^2 = o(|\alpha|^2)$$

that is for all $\varepsilon > 0$, there exists $r > 0$ such that for all α with $|\alpha|^2 \leq r$

$$\frac{|A^2(\hat{\varphi}(h)) - A^1(\hat{\varphi}(h)) + \eta \alpha^2|}{|\alpha|^2} \leq \varepsilon.$$

As $A^2(\hat{\varphi}(h)) - A^1(\hat{\varphi}(h))$ is positive and η is strictly positive:

$$\eta \leq \frac{A^2(\hat{\varphi}(h)) - A^1(\hat{\varphi}(h))}{\alpha^2} + \eta \leq \varepsilon$$

a contradiction (choose $\varepsilon = \eta/2$).

Now suppose that for all $0 \leq i < n$ $|\lambda_i^1| \geq |\lambda_i^2|$ and there exists $d \in \mathbb{R}^n$ such that $A^1(d) > A^2(d)$ that is there exists $\eta > 0$ such that $A^1(d) = A^2(d) + \eta$. Write $d = \sum_{i=0}^{n-1} d_i \hat{e}_i$ so that

$$\eta = A^1(d) - A^2(d) = \frac{1}{2} \sum_{i=0}^{n-1} (\lambda_i^1 - \lambda_i^2) d_i^2 + o(\|d\|_{\mathbb{R}^n}^2)$$

and for all $\varepsilon > 0$, there exists $r > 0$ such that for all d with $\|d\|_{\mathbb{R}^n}^2 \leq r$

$$\eta \leq \eta + \frac{\sum_{i=0}^{n-1} (\lambda_i^2 - \lambda_i^1) d_i^2}{\|d\|_{\mathbb{R}^n}^2} \leq \varepsilon$$

a contradiction.

As $\nabla^2 A^1(0_n) = -Q\Lambda Q^\top$, with q_{ij} being the elements of Q , this result implies that \succsim_1 is more ambiguity averse than \succsim_2 if and only if for all $0 \leq i < n$ and for all $0 \leq j < n$:

$$\frac{\partial^2 A^1}{\partial x_i \partial x_j}(0_n) = \sum_{k=0}^{n-1} \lambda_k^1 q_{ik} q_{jk} \leq \sum_{k=0}^{n-1} \lambda_k^2 q_{ik} q_{jk} = \frac{\partial^2 A^2}{\partial x_i \partial x_j}(0_n) \quad \square$$

A.3 Application to the choice of the Optimal Portfolio

Proof of Proposition 5.1 In our setting we have:

$$\text{sgn } \beta(\tilde{r}_1, \tilde{r}_2) = \text{sgn } \frac{\sigma_{12}}{\sigma_{11}} = \text{sgn } \sigma_{12}$$

and

$$\text{sgn } \alpha(\tilde{r}_1, \tilde{r}_2) = \text{sgn}(E_2 - \frac{\sigma_{12}}{\sigma_{11}} E_1) = \text{sgn}(E_2 \sigma_{11} - E_1 \sigma_{12})$$

As $\sigma_{11}\hat{\sigma}_{22} - \sigma_{12}^2 = \mathbf{var}(\tilde{r}_1) \mathbf{var}(\tilde{r}_2) - (\mathbf{cov}(\tilde{r}_1, \tilde{r}_2))^2 + \frac{\theta}{\gamma} \mathbf{var}(\tilde{r}_1) \mathbf{var}(\mathbb{A}\tilde{r}_2)$, the denominator in ψ_2^* is positive by the Cauchy-Schwarz inequality and the positivity of the variances. Therefore

$$\text{sgn } \psi_2^* = \text{sgn}(E_2\sigma_{11} - E_1\sigma_{12}) = \text{sgn } \alpha(\tilde{r}_1, \tilde{r}_2).$$

which proves the first relationship. Then the following derivatives:

$$\begin{aligned} \frac{\partial(\psi_1^*/\psi_2^*)}{\partial\theta} &= \frac{E_1}{E_2\sigma_{11} - E_1\sigma_{12}} \cdot \frac{\partial\hat{\sigma}_{22}}{\partial\theta} = \frac{E_1}{E_2\sigma_{11} - E_1\sigma_{12}} \cdot \frac{1}{\gamma} \mathbf{var}(\mathbb{A}\tilde{r}_2) \\ \frac{\partial\psi_2^*}{\partial\theta} &= -\frac{1}{\gamma} \frac{E_2\sigma_{11} - E_1\sigma_{12}}{(\sigma_{11}\hat{\sigma}_{22} - \sigma_{12}^2)^2} \cdot \sigma_{11} \frac{1}{\gamma} \mathbf{var}(\mathbb{A}\tilde{r}_2) \\ \frac{\partial\psi_1^*}{\partial\theta} &= \frac{1}{\gamma} \frac{E_1 \frac{1}{\gamma} \mathbf{var}(\mathbb{A}\tilde{r}_2)(\sigma_{11}\hat{\sigma}_{22} - \sigma_{12}^2) - \sigma_{11} \frac{1}{\gamma} \mathbf{var}(\mathbb{A}\tilde{r}_2)(E_1\hat{\sigma}_{22} - E_2\sigma_{12})}{(\sigma_{11}\hat{\sigma}_{22} - \sigma_{12}^2)^2} \\ &= \frac{1}{\gamma^2} \mathbf{var}(\mathbb{A}\tilde{r}_2) \sigma_{12} \frac{-E_1\sigma_{12} + \sigma_{11}E_2}{(\sigma_{11}\hat{\sigma}_{22} - \sigma_{12}^2)^2} \\ \frac{\partial(\psi_1^*/\psi_2^*)}{\partial\gamma} &= \frac{E_1}{E_2\sigma_{11} - E_1\sigma_{12}} \cdot \frac{\partial\hat{\sigma}_{22}}{\partial\gamma} = -\frac{E_1}{E_2\sigma_{11} - E_1\sigma_{12}} \cdot \frac{\theta}{\gamma^2} \mathbf{var}(\mathbb{A}\tilde{r}_2) \\ \frac{\partial\psi_2^*}{\partial\gamma} &= -\frac{E_2\sigma_{11} - E_1\sigma_{12}}{(\gamma\sigma_{11}\hat{\sigma}_{22} - \gamma\sigma_{12}^2)^2} (\sigma_{11}\sigma_{22} - \sigma_{12}^2) \end{aligned}$$

imply respectively

$$\begin{aligned} \text{sgn } \frac{\partial(\psi_1^*/\psi_2^*)}{\partial\theta} &= \text{sgn}(E_2\sigma_{11} - E_1\sigma_{12}) = \text{sgn } \alpha(\tilde{r}_1, \tilde{r}_2) \\ \text{sgn } \frac{\partial\psi_2^*}{\partial\theta} &= -\text{sgn}(E_2\sigma_{11} - E_1\sigma_{12}) = -\text{sgn } \alpha(\tilde{r}_1, \tilde{r}_2) \\ \text{sgn } \frac{\partial\psi_1^*}{\partial\theta} &= \text{sgn}(\sigma_{12}) \text{sgn}(E_2\sigma_{11} - E_1\sigma_{12}) = \text{sgn } \alpha(\tilde{r}_1, \tilde{r}_2) \beta(\tilde{r}_1, \tilde{r}_2) \\ \text{sgn } \frac{\partial(\psi_1^*/\psi_2^*)}{\partial\gamma} &= \text{sgn } \frac{\partial\psi_2^*}{\partial\gamma} = -\text{sgn}(E_2\sigma_{11} - E_1\sigma_{12}) = -\text{sgn } \alpha(\tilde{r}_1, \tilde{r}_2) \quad \square \end{aligned}$$

Proof of Proposition 5.2 Starting from the definitions, replacing and rearranging the terms lead to the results:

$$\begin{aligned} A &= E_1\hat{\sigma}_{22} - E_2\hat{\sigma}_{12} \\ &= E_1(\sigma_{C2}^2 + \sigma_{NC2}^2 + \frac{\theta}{\gamma}\sigma_{A2}^2) - E_2(\rho_C\sigma_{C1}\sigma_{C2} + \rho_{NC}\sigma_{NC1}\sigma_{NC2} + \frac{\theta}{\gamma}\rho_A\sigma_{A1}\sigma_{A2}) \\ &= (E_1\sigma_{C2}^2 - E_2\rho_C\sigma_{C1}\sigma_{C2}) + (E_1\sigma_{NC2}^2 - E_2\rho_{NC}\sigma_{NC1}\sigma_{NC2}) \\ &\quad + \frac{\theta}{\gamma} (E_1\sigma_{A2}^2 - E_2\rho_A\sigma_{A1}\sigma_{A2}) \end{aligned}$$

$$\begin{aligned} B &= E_2\hat{\sigma}_{11} - E_1\hat{\sigma}_{12} \\ &= E_2(\sigma_{C1}^2 + \sigma_{NC1}^2 + \frac{\theta}{\gamma}\sigma_{A1}^2) - E_1(\rho_C\sigma_{C1}\sigma_{C2} + \rho_{NC}\sigma_{NC1}\sigma_{NC2} + \frac{\theta}{\gamma}\rho_A\sigma_{A1}\sigma_{A2}) \\ &= (E_2\sigma_{C1}^2 - E_1\rho_C\sigma_{C1}\sigma_{C2}) + (E_2\sigma_{NC1}^2 - E_1\rho_{NC}\sigma_{NC1}\sigma_{NC2}) \\ &\quad + \frac{\theta}{\gamma} (E_2\sigma_{A1}^2 - E_1\rho_A\sigma_{A1}\sigma_{A2}) \end{aligned}$$

$$\begin{aligned}
C &= \hat{\sigma}_{11}\hat{\sigma}_{22} - \hat{\sigma}_{12}^2 \\
&= (\sigma_{C1}^2 + \sigma_{NC1}^2 + \frac{\theta}{\gamma}\sigma_{A1}^2)(\sigma_{C2}^2 + \sigma_{NC2}^2 + \frac{\theta}{\gamma}\sigma_{A2}^2) \\
&\quad - (\rho_C\sigma_{C1}\sigma_{C2} + \rho_{NC}\sigma_{NC1}\sigma_{NC2} + \frac{\theta}{\gamma}\rho_A\sigma_{A1}\sigma_{A2})^2 \\
&= \sigma_{C1}^2\sigma_{C2}^2 + \sigma_{C1}^2\sigma_{NC2}^2 + \frac{\theta}{\gamma}\sigma_{C1}^2\sigma_{A2}^2 + \sigma_{NC1}^2\sigma_{C2}^2 + \sigma_{NC1}^2\sigma_{NC2}^2 + \frac{\theta}{\gamma}\sigma_{NC1}^2\sigma_{A2}^2 \\
&\quad + \frac{\theta}{\gamma}\sigma_{A1}^2\sigma_{C2}^2 + \frac{\theta}{\gamma}\sigma_{A1}^2\sigma_{NC2}^2 + \frac{\theta^2}{\gamma^2}\sigma_{A1}^2\sigma_{A2}^2 - \rho_C^2\sigma_{C1}^2\sigma_{C2}^2 - \rho_{NC}^2\sigma_{NC1}^2\sigma_{NC2}^2 - \frac{\theta^2}{\gamma^2}\rho_A^2\sigma_{A1}^2\sigma_{A2}^2 \\
&\quad - 2\rho_C\sigma_{C1}\sigma_{C2}\rho_{NC}\sigma_{NC1}\sigma_{NC2} - 2\frac{\theta}{\gamma}\rho_C\sigma_{C1}\sigma_{C2}\rho_A\sigma_{A1}\sigma_{A2} - 2\frac{\theta}{\gamma}\rho_{NC}\sigma_{NC1}\sigma_{NC2}\rho_A\sigma_{A1}\sigma_{A2} \\
&= \sigma_{C1}^2\sigma_{C2}^2(1 - \rho_C^2) + \sigma_{NC1}^2\sigma_{NC2}^2(1 - \rho_{NC}^2) + \frac{\theta^2}{\gamma^2}\sigma_{A1}^2\sigma_{A2}^2(1 - \rho_A^2) \\
&\quad + \sigma_{C1}^2\sigma_{NC2}^2 + \sigma_{NC1}^2\sigma_{C2}^2 - 2\rho_C\sigma_{C1}\sigma_{C2}\rho_{NC}\sigma_{NC1}\sigma_{NC2} \\
&\quad + \frac{\theta}{\gamma}[\sigma_{C1}^2\sigma_{A2}^2 + \sigma_{A1}^2\sigma_{C2}^2 - 2\rho_C\sigma_{C1}\sigma_{C2}\rho_A\sigma_{A1}\sigma_{A2}] \\
&\quad + \frac{\theta}{\gamma}[\sigma_{A1}^2\sigma_{NC2}^2 + \sigma_{NC1}^2\sigma_{A2}^2 - 2\rho_{NC}\sigma_{NC1}\sigma_{NC2}\rho_A\sigma_{A1}\sigma_{A2}] \\
&= \sigma_{C1}^2\sigma_{C2}^2(1 - \rho_C^2) + \sigma_{NC1}^2\sigma_{NC2}^2(1 - \rho_{NC}^2) + \frac{\theta^2}{\gamma^2}\sigma_{A1}^2\sigma_{A2}^2(1 - \rho_A^2) \\
&\quad + (\sigma_{C1}\sigma_{NC2} - \sigma_{NC1}\sigma_{C2})^2 + 2\sigma_{C1}\sigma_{C2}\sigma_{NC1}\sigma_{NC2}(1 - \rho_C\rho_{NC}) \\
&\quad + \frac{\theta}{\gamma}[(\sigma_{C1}\sigma_{A2} - \sigma_{A1}\sigma_{C2})^2 + 2\sigma_{C1}\sigma_{C2}\sigma_{A1}\sigma_{A2}(1 - \rho_C\rho_A)] \\
&\quad + \frac{\theta}{\gamma}[(\sigma_{A1}\sigma_{NC2} - \sigma_{NC1}\sigma_{A2})^2 + 2\sigma_{NC1}\sigma_{NC2}\sigma_{A1}\sigma_{A2}(1 - \rho_{NC}\rho_A)] \quad \square
\end{aligned}$$

References

- ALIPRANTIS, C. D. AND K. C. BORDER (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer.
- ANSCOMBE, F. J. AND R. J. AUMANN (1963): "A Definition of Subjective Probability," *The Annals of Mathematical Statistics*, 34, 199–205.
- ARROW, K. J. (1965): "The theory of risk aversion," in *Aspects of the Theory of Risk Bearing*, Helsinki: Yrjö Jahnssonin Säätiö, reprinted in *Essays in the Theory of Risk Bearing* (1971), Chicago: Markham Publishing Co.
- BEWLEY, T. F. (2002): "Knightian decision theory. Part I," *Decisions in Economics and Finance*, 25, 79–110, (First version: 1986.).
- BOYLE, P. P., L. GARLAPPI, R. UPPAL, AND T. WANG (2012): "Keynes Meets Markowitz: The Trade-Off Between Familiarity and Diversification," *Management Science*, 58, 253–272.
- CHEN, Z. AND L. G. EPSTEIN (2002): "Ambiguity, Risk, and Asset Returns in Continuous Time," *Econometrica*, 70, 1403–1443.
- CONWAY, J. B. (1990): *A Course in Functional Analysis*, Springer.
- ELLSBERG, D. (1961): "Risk, Ambiguity, and the Savage Axioms," *The Quarterly Journal of Economics*, 75, 643–669.
- EPSTEIN, L. G. AND J. MIAO (2003): "A two-person dynamic equilibrium under ambiguity," *Journal of Economic Dynamics and Control*, 27, 1253–1288.

- FRENCH, K. R. AND J. M. POTERBA (1991): “Investor Diversification and International Equity Markets,” *The American Economic Review*, 81, 222–226.
- GARLAPPI, L., R. UPPAL, AND T. WANG (2007): “Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach,” *Review of Financial Studies*, 20, 41–81.
- GHIRARDATO, P., F. MACCHERONI, AND M. MARINACCI (2004): “Differentiating ambiguity and ambiguity attitude,” *Journal of Economic Theory*, 118, 133–173.
- GHIRARDATO, P. AND M. MARINACCI (2002): “Ambiguity Made Precise: A Comparative Foundation,” *Journal of Economic Theory*, 102, 251–289.
- GILBOA, I. AND D. SCHMEIDLER (1989): “Maxmin expected utility with non-unique prior,” *Journal of Mathematical Economics*, 18, 141–153.
- GOLLIER, C. (2011): “Portfolio Choices and Asset Prices: The Comparative Statics of Ambiguity Aversion,” *The Review of Economic Studies*, 78, 1329–1344.
- GOLUB, G. H. AND C. F. VAN LOAN (1996): *Matrix Computations*, John Hopkins Studies in the Mathematical Sciences, Baltimore and London: John Hopkins University Press, 3rd ed.
- IZHAKIAN, Y. AND S. BENNINGA (2011): “The Uncertainty Premium in an Ambiguous Economy,” *Quarterly Journal of Finance*, 01, 323–354.
- JEWITT, I. AND S. MUKERJI (2011): “Ordering Ambiguous Acts,” Economics Series Working Paper 553, University of Oxford, Department of Economics.
- KLIBANOFF, P., M. MARINACCI, AND S. MUKERJI (2005): “A Smooth Model of Decision Making under Ambiguity,” *Econometrica*, 73, 1849–1892.
- KREPS, D. (1988): *Notes On The Theory Of Choice*, Westview Press.
- MACCHERONI, F., M. MARINACCI, AND D. RUFFINO (Forthcoming): “Alpha as ambiguity: robust mean-variance portfolio analysis,” *Econometrica*.
- MACCHERONI, F., M. MARINACCI, AND A. RUSTICHINI (2006): “Ambiguity Aversion, Robustness, and the Variational Representation of Preferences,” *Econometrica*, 74, 1447–1498.
- MACCHERONI, F., M. MARINACCI, A. RUSTICHINI, AND M. TABOGA (2009): “Portfolio Selection with Monotone Mean-Variance Preferences,” *Mathematical Finance*, 19, 487–521.
- MARKOWITZ, H. M. (1952): “Portfolio Selection,” *The Journal of Finance*, 7, 77–91.
- MAS-COLELL, A., M. D. WHINSTON, AND J. R. GREEN (1995): *Microeconomic Theory*, New York: Oxford University Press.
- MEHRA, R. AND E. C. PRESCOTT (1985): “The equity premium: A puzzle,” *Journal of Monetary Economics*, 15, 145–161.
- NAU, R. F. (2003): “A Generalization of Pratt-Arrow Measure to Nonexpected-Utility Preferences and Inseparable Probability and Utility,” *Management Science*, 49, 1089–1104.
- (2006): “Uncertainty Aversion with Second-Order Utilities and Probabilities,” *Management Science*, 52, 136–145.

- PRATT, J. W. (1964): "Risk Aversion in the Small and in the Large," *Econometrica*, 32, 122–136.
- ROCKAFELLAR, R. T. (1970): *Convex Analysis*, no. 28 in Princeton Mathematical Series, Princeton, N.J: Princeton University Press.
- (2000): "Second-Order Convex Analysis," *Journal of Nonlinear and Convex Analysis*, 1, 1–16.
- SCHMEIDLER, D. (1989): "Subjective Probability and Expected Utility without Additivity," *Econometrica*, 57, 571–587.
- SEGAL, U. AND A. SPIVAK (1990): "First order versus second order risk aversion," *Journal of Economic Theory*, 51, 111–125.
- SINISCALCHI, M. (2009): "Vector Expected Utility and Attitudes Toward Variation," *Econometrica*, 77, 801–855.
- SKIADAS, C. (2008): "Smooth Ambiguity Aversion Toward Small Risks and Continuous-Time Recursive Utility," *SSRN eLibrary*.
- TABOGA, M. (2005): "Portfolio selection with two-stage preferences," *Finance Research Letters*, 2, 152–164.
- TOBIN, J. (1958): "Liquidity Preference as Behavior Towards Risk," *The Review of Economic Studies*, 25, 65–86.